DOCTORAL THESIS

WAVELETS, DIFFUSION EQUATIONS AND THEIR APPLICATIONS

AUTHOR

SANTOSH KUMAR

SUPERVISOR

DR. MD. KALIMUDDIN AHMAD

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

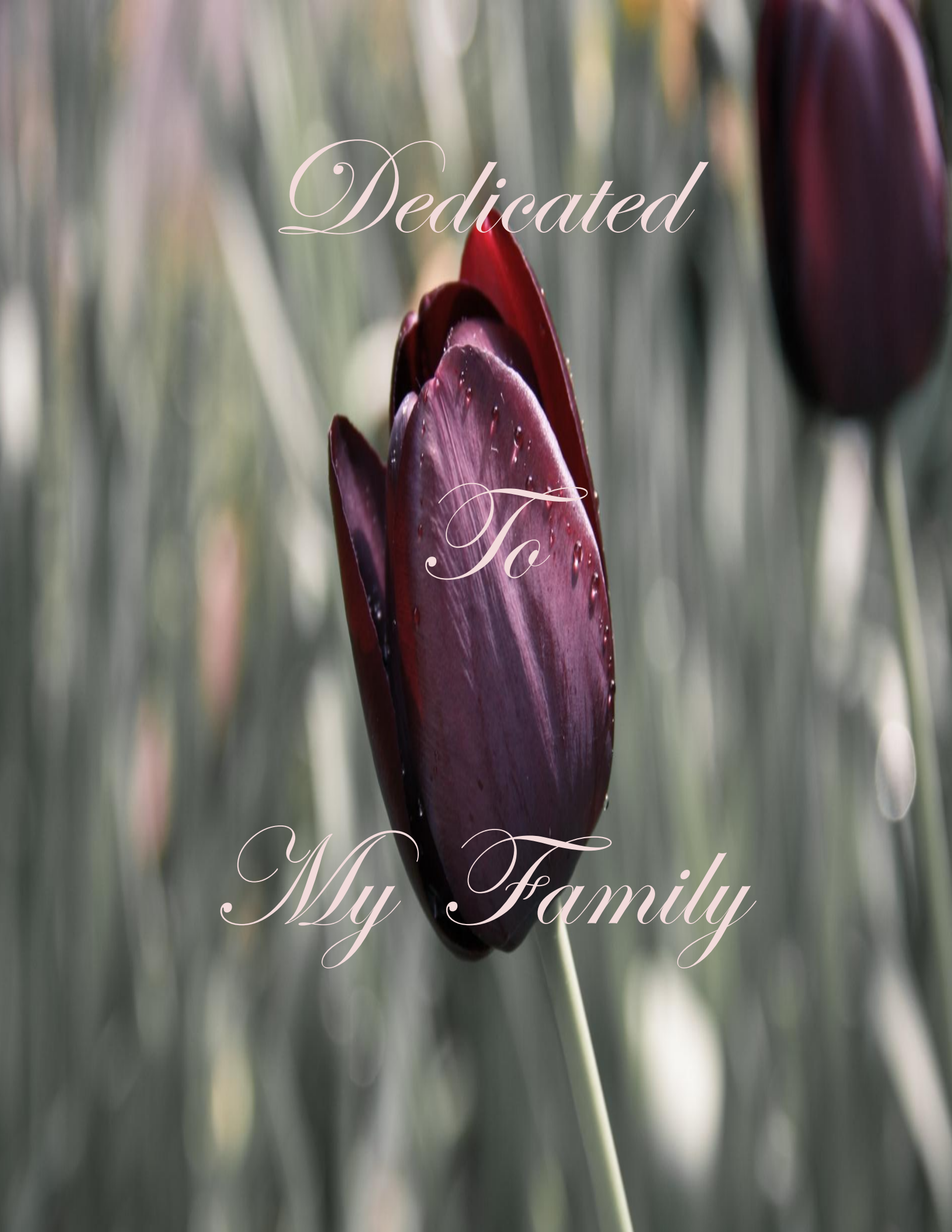
In

Mathematics

at

Department of Mathematics
Aligarh Muslim University
Aligarh-202002, India

2016


Dedicated
To
My Family

Dr. M. K. Ahmad
M.Phil, M.Sc. Nat. (Germany), Ph.D.
DAAD Fellow (Germany) 1991-93 & 1996-98
Associate Professor

Department of Mathematics
Aligarh Muslim University, Aligarh
U.P. - 202 002 (INDIA)
✆ : +91(0)571-2704736 (R)
09456973184 (M)
Fax: +91(0) 571-2700528

Certificate

*This is to certify that the thesis entitled "**Wavelets, diffusion equations and their applications**" is the research work of **Mr. Santosh Kumar** carried out under my supervision and guidance. He has fulfilled the prescribed conditions given in the statues and ordinances of Aligarh Muslim University, Aligarh.*

I further certify that the work of this thesis either partially or fully has not been submitted to any other University or Institution for the award of any degree.

(Dr. Md. Kalimuddin Ahmad)
Supervisor

Specializations : Image Processing, Wavelets, Mathematical Modeling and V. Inequalities
Residence : A-3, Naaz Residency, New Sir Syed Nagar, Aligarh, U.P. – 202 002
E-mails : ahmad_kalimuddin@yahoo.co.in; mk.ahmad.mm@amu.ac.in

DEPARTMENT OF MATHEMATICS
ALIGARH MUSLIM UNIVERSITY
ALIGARH - 202002, INDIA

Candidate's Declaration

I, Santosh Kumar, Department of Mathematics, certify that the work embodied in this Ph.D. thesis is my own bonafide work carried out by me under the supervision of Dr. M.K. Ahmad at Aligarh Muslim University, Aligarh. The matter embodied in this Ph.D. thesis has not been submitted for the award of any other degree.

I declare that I have faithfully acknowledged, given credit to and referred to the research workers wherever their work has been cited in the text and the body of the thesis. I further certify that I have not willfully lifted up some other's work, paragraph, text, data, result etc. reported in the journals, books, magazines, reports, dissertations, thesis etc. or available at web-sites and included them in this Ph.D. thesis and cited as my own work.

Date:

(Santosh Kumar)

.....

Certificate from the Supervisor

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Signature of the Supervisor:...……..

Name and Designation: Dr. M.K. Ahmad
Associate Professor

Department: Mathematics

(Signature of the Chairman of the Department with seal)

DEPARTMENT OF MATHEMATICS
ALIGARH MUSLIM UNIVERSITY
ALIGARH - 202002, INDIA

Course/Comprehensive Examination/Pre-Submission Seminar Completion Certificate

This is to certify that ***Mr. Santosh Kumar***, Department of Mathematics, has satisfactorily completed the course work/comprehensive examination and pre-submission seminar requirement which is part of his Ph.D. programme as required under the UGC (Minimum Standards and Procedure for Award of M.Phil./Ph.D. Degree) Regulation, 2009.

Date: **(Signature of the Chairman**

of the Department)

Copyright Transfer Certificate

Title of the Thesis: Wavelets, diffusion equations and their Applications

Candidate's Name: Santosh Kumar

Copyright Transfer

The undersigned hereby assigns to Aligarh Muslim University, Aligarh copyright that may exist in and for the above thesis submitted for the award of the Ph.D. degree.

Signature of the candidate

Note: However, the author may reproduce or authorize others to reproduce material extracted verbatim from the thesis or derivative of the thesis for author's personal use provide that the source and the University's copyright notice are indicated.

Table of Contents

Acknowledgement

All praises and thanks be to Almighty **'GOD'**, the most beneficent, the most merciful, Who bestowed upon me the courage, patience and strength to embark upon this work and carry it to its completion.

This thesis is the end of my journey in obtaining my Ph.D. This thesis has been kept on track and been seen through to completion with the support and encouragement of numerous people including my well wishers, friends and colleagues.

I have immense pleasure to express my gratefulness and indebtedness to my supervisor **Dr. Md. Kalimuddin Ahmad**, Associate Professor, Department of Mathematics, Aligarh Muslim University, Aligarh, for boosting up my panache and prodding me to accomplish this task. His continuous guidance and invaluable suggestions instilled me with morale needed to complete the work. The critical comments, he rendered during the discussion have gone a long way in my understanding and presentation of the contents of this thesis.

I wish to thank Prof. Mursaleen, Chairman, Department of Mathematics, Aligarh Muslim University, Aligarh for providing me departmental facilities throughout the development stages of this thesis and encouragement in pursuit of course of study. I also avail this opportunity to express my thanks to all my teachers for their encouragement and inspiration which in turn helped me towards perfection.

I would like to express my gratitude to all my seniors especially Dr. Javid Iqbal, Dr. Sana Khan, Dr. Asif Khan, for their generous help, suggestions and co-operation. It is my pleasure to express my deep sense of appreciation to my friends and colleagues especially Mr. Mohd. Sarfaraz, Dr. Khursheed Jamal Ansari, Mr. Mijanur Rahamana, Mr. Rehan Ali who always supported my aspiration with ease and love at various stages of this work. Sincerely thanks are due to my juniors specially to Mr. Rajat Vaish, Mr. Aadil Hussain Dar for their constant involvement and all kind of assistance with motivation, encouragement and guidance. They had always been there with a helping hand throughout the course of my work.

My deepest sense of gratitude and indebtedness goes to my mother **Mrs. Ram Bati**, father **Late Mr. Hakim Singh**, wife **Mrs. Sangeeta Yadav** and brother **Mr. Durgesh Kumar** for their unflagging love, continuous and constant encouragement throughout my life and above all for their moral support especially in the period

of doubt and despair. I always fall short of words and felt impossible to describe their support in words. Their love and prayers are always my confidence and will boost my progress. They have always given me strength and wisdom to be sincere in my work, for setting high moral standards and supporting me through their hard work, love and affection. Their support and encouragement was in the end which makes this dissertation possible. Thanks for giving me solid roots from which I can grow, for being a constant and active presence in my life, for accepting my dreams even if they were different from theirs and for doing their best in helping me to realize them. I owe them everything and wish I could show them just how much I love and appreciate them.

Finally, I express my indebtedness to the seminar library staff for providing me books, journals and other facilities. I express my indebtedness to my glorious and esteemed institution Aligarh Muslim University, Aligarh, my homage to its founder Sir Syed Ahmad Khan and U.G.C, for providing me financial assistance in the form of Junior Research Fellowship (UGC-JRF) during my Ph.D programme.

At last, I thank to all who genuinely offered wish and help. I appreciate everyone whom I have missed to remember and have contributed towards completion of the work.

Dated: **(Santosh Kumar)**

Preface

Partial differential equations (PDEs) have led to an entire new field in image processing and computer vision. Hundreds of publications have appeared in the last decade, and PDE-based methods have played a central role at several conferences and workshops. The success of these techniques is not really surprising, since PDEs have proved their usefulness in areas such as physics and engineering sciences for a very long time. In image processing and computer vision, they offer several advantages:

- Deep mathematical results with respect to well-posedness are available, such that stable algorithms can be found. PDE-based methods are one of the mathematically best-founded techniques in image processing.

- They allow a reinterpretation of several classical methods under a novel unifying framework. This includes many well-known techniques such as Gaussian convolution, median filtering, dilation or erosion.

- This understanding has also led to the discovery of new methods. They can offer more invariances than classical techniques, or describe novel ways of shape simplification, structure preserving filtering, and enhancement of coherent line-like structures.

- The PDE formulation is genuinely continuous. Thus, their approximations aim to be independent of the underlying grid and may reveal good rotational invariance.

PDE-based image processing techniques are mainly used for smoothing and restoration purposes. Many evolution equations for restoring images can be derived as gradient descent methods for minimizing a suitable energy functional, and the restored image is given by the steady-state of this process. Typical PDE techniques for image smoothing regard the original image as initial state of a parabolic (diffusion-like) process, and extract filtered versions from its temporal evolution. The whole evolution can be regarded as a so-called scale-space, an embedding of the original image into a family of subsequently simpler, more global representations of it. Since this introduces a hierarchy into the image structures, one can use a scale-space representation for extracting semantically important information.

An interesting class of parabolic equations which may bridge the gap between scale-space and restoration ideas: nonlinear diffusion filters. Methods of this type have been proposed for the first time by Perona and Malik in 1990 [56]. In order to smooth an image and to simultaneously enhance important features such as edges, they apply a diffusion process whose diffusivity is steered by derivatives of the evolving image. These filters are difficult to analyse mathematically, as they may act locally like a backward diffusion process. This gives rise to well-posedness questions. On the other hand, nonlinear diffusion filters are frequently applied with very impressive results; so there appears the need for a theoretical foundation.

The wavelets, on the other hand, is an excellent tool in the field of data compression. It's correspondence with the diffusion filters has been studied by many researchers, see especially the papers of Mrázek et al. [50, 51] and and Welk et al. [81].

The present thesis comprises of six chapters and each chapter is divided into sections. The definitions, examples, remarks, theorems etc. have been specified with the double decimal numbers. The first figure denotes the number of chapter, second represents the section in the chapter and the third points out the number of the definition, the example and the theorem as the case may be in a particular chapter.

Chapter 1 contains preliminary notions, basic definitions, examples and relevant well known results related to our study which are required for the development of the subject in subsequent chapters.

Chapter 2 is divided into six sections. The first section is introductory; the second section deals with the concept of new nonlinear diffusion model for image restoration; the third section contains the result which guarantees the existence, uniqueness and stability of viscosity solution of our model defined in previous section. In fourth section, we have defined a discrete scheme for incorporating numerical experiments which are given in fifth section. The last section deals with the conclusion.

Chapter 3 is divided into seven sections. The first section is introductory. In second section, we have defined a new nonlinear anisotropic diffusion model for image denoising. In section three, we have proved a result which establishes the existence and uniqueness of weak solutions obtained by our new model defined in previous section. In the fourth section, we discuss about some properties of weak solution. In fifth section, we have defined a convergent iterative scheme and with the help of which we have

exhorted some numerical experiments in sixth section. The last section concludes the chapter.

Chapter 4 is divided into nine sections. The first section is introductory. Second and third sections deal with the physical background and mathematical consideration of anisotropic diffusion model for image denoising respectively; fifth and sixth sections deal with the study of concept of new time dependent models for 2D and 1D respectively. We have compared our models with the existing models by means of standard metrices in fourth, seventh and eighth sections. The last section deals with the conclusion.

Chapter 5 is divided into five sections. The first section is introductory. In second section, we have defined a new PDE based deconvolution model. In third section, we have used two different types of diffusivity in our model defined in previous section. Numerical experiments has been done in fourth section. The last section deals with the conclusion.

Chapter 6 is devoted to the study of reconstruction of wavelet coefficients using nonlinear diffusion reaction equation. The first section of this chapter is introductory; the second section deals with the study of discrete wavelet transform. In the third section, we have defined a new diffusion equation based on wavelet coefficients. In fourth section, we have defined a discrete scheme for incorporating numerical experiments which are given in fifth section. The last section deals with the conclusion.

In the end, a comprehensive bibliography is included.

Chapter 1

Preliminaries

1.1 Introduction

The fundamental idea behind wavelets is to analyze according to scale. Indeed, some researchers in the wavelet field feel that, by using wavelets, one is adopting a whole new mind set or perspective in processing data. Wavelets are functions that satisfy certain mathematical requirements and are used in representing data or other functions. This idea is not new. Approximation using superposition of functions has existed since the early 1800's, when *Joseph Fourier* discovered that he could superpose sines and cosines to represent other functions. However, in wavelet analysis, the scale that we use to look at data plays a special role. Wavelet algorithms process data at different scales or resolutions. If we look at a signal with a large "window", we would notice gross features. Similarly, if we look at a signal with a small "window", we would notice small features. This makes wavelets interesting and useful. For many decades, scientists have wanted more appropriate functions than the sines and cosines which comprise the bases of Fourier analysis, to approximate choppy signals [23]. By their definition, these functions are non-local (and stretch out to infinity). They therefore do a very poor job in approximating sharp spikes. But with wavelet analysis, we can use approximating functions that are contained neatly in finite domains. Wavelets are well-suited for approximating data with sharp discontinuities. The wavelet analysis procedure is to adopt a wavelet prototype function, called an "*analyzing wavelet*" or "*mother wavelet*". Temporal analysis is performed with a contracted, high-frequency version of the prototype wavelet, while frequency analysis is performed with a dilated, low-frequency version of the same wavelet. Because the original signal or function

can be represented in terms of a wavelet expansion (using coefficients in a linear combination of the wavelet functions), data operations can be performed using just the corresponding wavelet coefficients. And if you further choose the best wavelets adapted to your data, or truncate the coefficients below a threshold, your data is sparsely represented. This sparse coding makes wavelets an excellent tool in the field of data compression. Other applied fields that are making use of wavelets include astronomy, acoustics, nuclear engineering, sub-band coding, signal and image processing, neurophysiology, music, magnetic resonance imaging, speech discrimination, optics, fractals, turbulence, earthquake-prediction, radar, human vision, and pure mathematics applications such as solving partial differential equations.

The 1930S

In the 1930s, several groups working independently researched the representation of functions using $scale-varying\ basis\ functions$. Understanding the concepts of basis functions and scale-varying basis functions is key to understanding wavelets. By using a scale-varying basis function called the Haar basis function, Paul Levy, a 1930s physicist, investigated Brownian motion, a type of random signal [48]. He found the Haar basis function superior to the Fourier basis functions for studying small complicated details in the Brownian motion. Another 1930s research effort by Littlewood, Paley, and Stein involved computing the energy of a function $f(x)$:

$$energy = \frac{1}{2}\int_0^{2\pi} |f(x)|^2 dx.$$

The computation produced different results if the energy was concentrated around a few points or distributed over a larger interval. This result disturbed the scientists because it indicated that energy might not be conserved. The researchers discovered a function that can vary in scale and can conserve energy when computing the functional energy. Their work provided $David\ Marr$ with an effective algorithm for numerical image processing using wavelets in the early $1980s$.

1960-1980

Between 1960 and 1980, the mathematicians $Guido\ Weiss$ and $Ronald\ R.\ Coifman$ studied the simplest elements of a function space, called $atoms$, with the goal of finding

the atoms for a common function and finding the "assembly rules" that allow the reconstruction of all the elements of the function space using these atoms. In 1980, *Grossman* and *Morlet*, a physicist and an engineer, broadly defined wavelets in the context of quantum physics. These two researchers provided a way of thinking for wavelets based on physical intuition.

Post-1980

In 1985, *Stephane Mallat* gave wavelets an additional jump-start through his work in digital signal processing. He discovered some relationships between quadrature mirror filters, pyramid algorithms, and orthonormal wavelet bases. Inspired in part by these results, *Y. Meyer* constructed the first non-trivial wavelets. Unlike the Haar wavelets, the Meyer wavelets are continuously differentiable; however they do not have compact support. A couple of years later, *Ingrid Daubechies* used Mallat's work to construct a set of wavelet orthonormal basis functions that are perhaps the most elegant, and have become the cornerstone of wavelet applications today.

This Chapter is introductory. We have given some basic definitions and concepts which are used in the subsequent chapters.

Definition 1.1.1. *Let $L^2(0, 2\pi)$ denote the collection of all measurable functions f (we can assume f a piecewise continuous function) defined on the interval $(0, 2\pi)$ with*

$$\int_0^{2\pi} | f(x) |^2 \, dx < \infty.$$

The collection $L^2(0, 2\pi)$ is often called the space of 2π-periodic square-integrable functions.

Definition 1.1.2. *The space $L^2(\mathbb{R})$ of measurable functions f on real line $\mathbb{R}$ is defined as $\{f : \mathbb{R} \to \mathbb{R} : \int_{-\infty}^{\infty} | f(x) |^2 \, dx < \infty\}$.*
The norm of $f(x)$ in $L^2(\mathbb{R})$ is defined as

$$\|f\|_2^2 = \int_{-\infty}^{\infty} | f(x) |^2 \, dx,$$

or,

$$\|f\|_2 = \langle f, f\rangle^{\frac{1}{2}}.$$

Definition 1.1.3. *For a pair of functions f and g belonging to $L^2(\mathbb{R})$, the inner product of f and g is defined as*

$$\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)\overline{g(x)}dx,$$

where $\overline{g(x)}$ is the complex conjugate of $g(x)$.

Definition 1.1.4. *$l^2(\mathbb{Z})$ is the vector space of square summable sequences, i.e., $\{c_n\} \in l^2(\mathbb{Z})$ iff $\sum_{n=-\infty}^{\infty} \mid c_n \mid^2 < \infty$.*

Definition 1.1.5. *The Fourier transform of $f(x) \in L^2(\mathbb{R})$ is written as $\widehat{f}(\omega)$ and is defined by*

$$\widehat{f}(\omega) = \int_{-\infty}^{\infty} f(x)e^{-i\omega x}dx.$$

Definition 1.1.6. *If $f \in L^1(\mathbb{R})$ and its Fourier transform $\hat{f} \in L^1(\mathbb{R})$, then the inverse Fourier transform of $\hat{f}(\omega)$ is defined by*

$$f(x) = \mathcal{F}^{-1}\{\hat{f}(\omega)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega x}\hat{f}(\omega)d\omega.$$

In general, f can be reconstructed from $\hat{f}$ at each point $x \in \mathbb{R}$ where f is continuous.

Definition 1.1.7. *The convolution of two function $f, g \in L^1(\mathbb{R})$ is defined by*

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x-y)g(y)dy, \ \ \forall\, x \in \mathbb{R}.$$

In case of two dimension, the convolution of functions $f, g \in L^1(\mathbb{R}^2)$ is defined as follows:

$$(f(x,y) * g(x,y)) = \int_{\tau=-\infty}^{\infty}\int_{\tau=-\infty}^{\infty} f(\tau_1,\tau_2)g(x-\tau_1, y-\tau_2)d\tau_1 d\tau_2, \ \ \forall\, (x,y) \in \mathbb{R}^2.$$

In 2D-discrete space:

$$c[m,n] = a[m,n] * b[m,n] = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a[j,k]b[m-j,n-k].$$

1.1.1 Windowed Fourier transform

If $f(t)$ is a non-periodic signal, the summation of the periodic functions, sine and cosine, does not accurately represent the signal. You could artificially extend the signal to make it periodic but it would require additional continuity at the endpoints. The windowed Fourier transform (WFT) is one solution to the problem of better representing the non-periodic signal. The WFT can be used to give information about signals simultaneously in the time domain and in the frequency domain. With the WFT, the input signal $f(t)$ is chopped up into sections, and each section is analyzed for its frequency content separately. If the signal has sharp transitions, we window the input data so that the sections converge to zero at the end points [37]. This windowing is accomplished via a weight function that places less emphasis near the interval's end points than in the middle. The effect of the window is to localize the signal in time.

To approximate a function by samples, and to approximate the Fourier integral by the discrete Fourier transform, requires applying a matrix whose order is the number sample points n. Since multiplying an $n \times n$ matrix by a vector costs on the order of n^2 arithmetic operations, the problem gets quickly worse as the number of sample points increases. However, if the samples are uniformly spaced, then the Fourier matrix can be factored into a product of just a few sparse matrices, and the resulting factors can be applied to a vector in a total of order $n \log n$ arithmetic operations. This is the so-called fast Fourier transform or FFT [58].

The deficiency of the formula of F.T. in time frequency analysis was observed by D. Gabor in 1946. He introduced a "time localization" function $g(t-u)$, where the parameters u is used to translate the window in order to cover the whole time domain, for extracting local information of the signal.

Definition 1.1.8. [21] *At a position u and for a frequency ω, the windowed F.T. or a*

Gabor transform of a function $f(x) \in L^2(\mathbb{R})$ is defined by

$$Gf(\omega, u) = \int_{-\infty}^{\infty} e^{-i\omega x} g(x-u) f(x) dx.$$

Definition 1.1.9. *A function $\psi(x) \in L^2(\mathbb{R})$ is called a wavelet or mother wavelet if the following condition is satisfied:*

$$C_\psi := \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty, \tag{1.1.1}$$

where, $\hat{\psi}$ is the Fourier transform of ψ.

Remark 1.1.1. *(i) Condition (1.1.1) implies that*

$$\int_{-\infty}^{\infty} \psi(t) dt = 0. \tag{1.1.2}$$

Very often a function $\psi \in L^2(\mathbb{R})$ satisfying (1.1.2) is called a wavelet. However, in order to get analogue of inversion formula for Fourier transform we require stronger condition (1.1.1).
(ii) For $\psi \in L^2(\mathbb{R})$ satisfying $t\psi \in L^1(\mathbb{R})$, i.e., $\int_{-\infty}^{\infty} t\psi(t) dt < \infty$, condition (1.1.1) is equivalent to

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad \textit{resp.} \quad \hat{\psi}(0) = 0.$$

Lemma 1.1.1. *Let ϕ be a nonzero n-times $(n \geq 1)$ differentiable function such that $\phi^{(n)} \in L^2(\mathbb{R})$. Then $\psi = \phi^{(n)}(x)$ is a wavelet.*

Definition 1.1.10. *Relative to every basic wavelet ψ, the integral wavelet transform (IWT) on $L^2(\mathbb{R})$ is defined by*

$$(W_\psi f)(b, a) := |a|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t) \psi\left(\overline{\frac{t-b}{a}}\right) dt, \quad f \in L^2(\mathbb{R}). \tag{1.1.3}$$

Remark 1.1.2. *If we consider $\psi_{b,a}(t)$ as a family of functions given by*

$$\psi_{b,a}(t) = |a|^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right), \quad a > 0, \quad b \in \mathbb{R}$$

where ψ is a fixed function, often called mother wavelet. Then (1.1.3) can be written as

$$(W_\psi f)(b,a) = \langle f, \psi_{b,a} \rangle.$$

Theorem 1.1.1. *Let ψ be a wavelet and ϕ a bounded integrable function, then the convolution function $\psi * \phi$ is a wavelet.*

Theorem 1.1.2. *(Parseval's formula for wavelet transforms) Let $\psi \in L^2(\mathbb{R})$ be a wavelet. Then, for any $f, g \in L^2(\mathbb{R})$, the following formula hold:*

$$\langle f, g \rangle = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (W_\psi f)(b,a) \overline{(W_\psi g)(b,a)} \, \frac{dbda}{a^2}.$$

Theorem 1.1.3. *(Inversion formula) Let $\psi \in L^2(\mathbb{R})$ be a wavelet. Then, for any $f \in L^2(\mathbb{R})$, the following formula hold:*

$$f(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (W_\psi f)(b,a) |a|^{-\frac{1}{2}} \psi \left(\frac{t-b}{a} \right) \, \frac{dbda}{a^2}.$$

Definition 1.1.11. [11] *A wavelet $\psi \in L^2(\mathbb{R})$ is said to have n vanishing moments if*

$$\int_{-\infty}^{\infty} t^k \psi(t) \, dt \; = \; 0, \quad \textit{for} \;\; 0 \leq k < n.$$

1.2 Examples of wavelets

Example 1.2.1. *(Haar wavelet) The function defined below is called the Haar function*

$$\psi(x) = \begin{cases} 1, & 0 \leq x < \frac{1}{2} \\ -1, & \frac{1}{2} \leq x < 1 \\ 0, & \textit{otherwise.} \end{cases}$$

ψ has compact support, i.e. ψ is zero outside $[0,1)$. Since $\psi \in L^2(\mathbb{R})$ and also satisfied the condition (1.1.1), therefore the Haar function is a wavelet.

Example 1.2.2. *(Mexican hat wavelet) The function ψ defined as*

$$\psi(x) = (1 - x^2) e^{-x^2/2}$$

is known as the Mexican hat function. It is a wavelet by Lemma 1.1.1. Mexican hat wavelet is the second derivative of Gaussian function $-e^{-x^2/2}$.

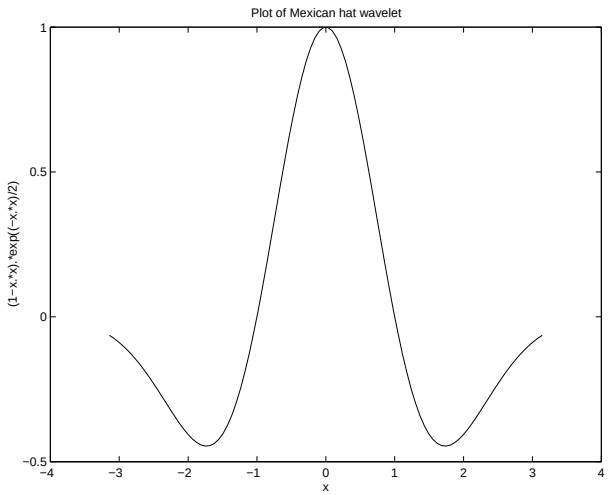


Figure 1.2.1: Mexican hat wavelet

Example 1.2.3. *(Morlet or Gabor wavelet) A function defined as*

$$\psi(x) = (e^{i\alpha t} - e^{-\alpha^2/2})e^{-t^2/2}$$

is called Morlet or Gabor function. Since $\psi \in L^2(\mathbb{R})$, $x\psi(x) \in L^1(\mathbb{R})$, *by Remark 1.1.1(ii),* ψ *is a wavelet and is known as Morlet or Gabor wavelet.*

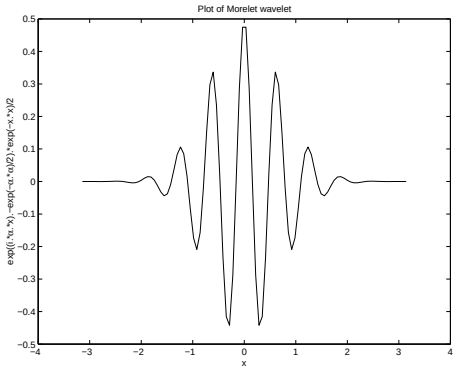


Figure 1.2.2: Morlet wavelet

1.3 An overview of diffusion process

The main aim of this section is to provide a physical background of diffusion process. As PDE-based methods appear in a large variety of image processing and computer vision areas ranging from shape-from-shading and histogram modification to optic flow and stereo vision.

Most people have an intuitive impression of diffusion as a physical process that equilibrates concentration differences without creating or destroying mass. This physical observation can be easily cast in a mathematical formulation. The equilibration property is expressed by *Fick's law*:

$$j = -D \cdot \nabla u. \tag{1.3.1}$$

This equation states that a concentration gradient ∇u causes a flux j which aims to compensate for this gradient. The relation between ∇u and j is described by the diffusion tensor D, a positive definite symmetric matrix. The case where j and ∇u are parallel is called "isotropic". Then we may replace the diffusion tensor by a positive scalar-valued diffusivity g. In the general anisotropic case, j and ∇u are not parallel.

The observation that diffusion does only transport mass without destroying it or creating new mass is expressed by the *continuity equation*

$$\frac{\partial u}{\partial t} = -divj, \tag{1.3.2}$$

where t denotes the time.

If we plug in Ficks law into the continuity equation we end up with the *diffusion equation*

$$\frac{\partial u}{\partial t} = div(D \cdot \nabla u). \tag{1.3.3}$$

This equation appears in many physical transport processes. In the context of heat transfer it is called *heat equation*. In image processing we may identify the concentration with the grey value at a certain location. If the diffusion tensor is constant over the whole image domain, one speaks of *homogeneous diffusion*, and a space-dependent filtering is called *inhomogeneous*. Often the diffusion tensor is a function of the differential structure of the evolving image itself. Such a feedback leads to *nonlinear diffusion filters*. Diffusion which does not depend on the evolving image is called *linear*.

1.4 Linear diffusion filter

The simplest and best investigated PDE in image processing is the parabolic linear diffusion equation of the form

$$\frac{\partial u}{\partial t} = \Delta u(x,t), \quad (x,t) \in \mathbb{R}^2 \times (0,\infty) \tag{1.4.1}$$

with initial condition $u(x,0) = f(x)$ for any $x \in \mathbb{R}^2$. In fact, f is primarily defined only on the domain $\Omega \subset \mathbb{R}^2$, but by symmetry and then periodicity we can extend it to $\mathbb{R}^2$. This method of extension is typical in image processing.

The underlying idea to apply equation (1.4.1) in image processing comes from the early work of Koenderink [39], who noticed that the convolution of the image f with the Gaussian kernel, defined by

$$G_\sigma(x,y) := \frac{1}{2\pi\sigma^2}\exp\left(-\frac{|x-y|^2}{2\sigma^2}\right) \tag{1.4.2}$$

is equivalent to the solution u of problem (1.4.1) for $t = \frac{1}{2}\sigma^2$, that is,

$$u(x,t) = (G_{\sqrt{2t}} * f)(x) = \int_{\mathbb{R}^2} G_{\sqrt{2t}}(x,y) f(y) dy.$$

The above formula gives the correspondence between the time t and the scale parameter σ of the Gaussian kernel G_σ.

The linear diffusion filter has one serious disadvantage. As a matter of fact, it

smoothens an image but at the same time blurs important features such as edges, making it difficult to identify on the next stage of image analysis, namely segmentation, see the figure 1.8.1. To overcome this problem one should consider a nonlinear filter, adapted to the local image structure.

1.5 Nonlinear isotropic diffusion filter

1.5.1 The Perona-Malik model

For the first time a nonlinear diffusion filter was introduced by Perona and Malik [56]. They proposed to replace (1.4.1) by a nonlinear diffusion equation with homogeneous Neumann conditions on the boundary $\partial\Omega$ and solve the following problem in order to obtain the smoothened version of the initial image f

$$\begin{cases} \frac{\partial u}{\partial t} = \nabla \cdot (d(|\nabla u|^2)\nabla u), \ (x,t) \in \Omega \times (0,T], \\ \\ u(x,0) = f, \quad x \in \Omega, \\ \\ \frac{\partial u}{\partial n} = 0, \quad (x,t) \in \Omega \times (0,T]. \end{cases} \tag{1.5.1}$$

In the first equation, the diffusivity d is positive, monotonically decreasing function, defined in a way, such that the smoothing of image is conditional and depends on its structure. If $|\nabla u|^2$ is large, then the diffusion is low and therefore exact location of edges is kept. If $|\nabla u|^2$ is small, then the diffusion tends to smooth around x. Of course, there exist several possible choices for diffusivity d. As an example, consider the following definition

$$d(s) := \frac{1}{1 + s/\mu}, \tag{1.5.2}$$

where the parameter $\mu > 0$ plays the role of threshold.

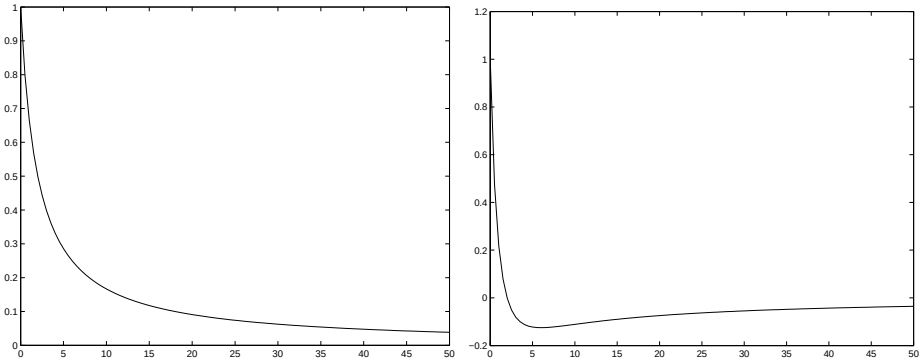


Figure 1.5.1: (a) Left: Plot of the diffusivity $d(s) = \frac{1}{1+s/\mu}$ with threshold $\mu = 2$ and (b) Right: Plot of the function $b(s) = d(s) + 2sd'(s)$.

However, the Perona-Malik model has several practical and theoretical difficulties. If the image is noisy, then the noise would introduce very large oscillations of the gradient ∇u. Therefore, the adaptive smoothing introduced by the model (1.5.1) would not give good results, since all these noise edges will be kept. The second difficulty arises from the fact that we obtain a backward diffusion equation for $|\nabla u|^2 > \mu$, which is a classical example of the ill-posed problem. In practice, it implies that very similar images can give divergent solutions and therefore different edges.

1.5.2 Regularization of the Perona-Malik model

One way to deal with an ill-posed problem like (1.5.1) is to introduce regularization which would make the problem well-posed. Then, by reducing the amount of regularization and observing the behavior of the solution of the regularized problem, we can obtain information about the initial one. In the first time, the method to regularize the Perona-Malik problem was proposed by Catté et al. [13]. They suggested to replace the gradient ∇u in the diffusivity $d(|\nabla u|^2)$ by its estimate $\nabla u_\sigma := \nabla G_\sigma * u$, where G_σ is the Gaussian kernel as defined in (1.4.2). They have also proven that this slight change is sufficient to ensure existence and uniqueness of the solution to the problem (1.5.1). This result gives the following theorem.

Theorem 1.5.1. *Let $d : \mathbb{R}_+ \cup \{0\} \to \mathbb{R}_+$ be smooth, decreasing with $d(0) = 1$, $\lim_{s\to\infty} d(s) = 0$. If $f \in L^2(\Omega)$, then there exists an unique function $u(x,t) \in C([0,T], L^2(\Omega)) \cap L^2((0,T), H^1(\Omega))$ satisfying in the distributional sense*

$$\begin{cases} \frac{\partial u}{\partial t} = \nabla \cdot (d(|\nabla u_\sigma|^2)\nabla u), \ (x,t) \in \Omega \times (0,T], \\ \\ u(x,0) = f, \quad x \in \Omega, \\ \\ \frac{\partial u}{\partial n} = 0, \quad (x,t) \in \partial\Omega \times (0,T]. \end{cases} \tag{1.5.3}$$

Moreover, $|u|_{L^\infty((0,T),L^2(\Omega))} \leq |f|_{L^2(\Omega)}$ *and* $u \in C^\infty((0,T) \times \overline{\Omega})$.

1.6 Nonlinear anisotropic diffusion filter

1.6.1 The Weickert model

Despite the advantages of the isotropic diffusion filter, there is still one imperfection: when the diffusion process is stopped near the boundary of an object, it preserves the edges but also leaves a noise at these positions. To avoid this effect, Weickert [78] suggested to modify the diffusion operator so that it diffuses more in direction parallel to edges and less in the perpendicular one. In order to filter an image, he proposed to consider the following problem

$$\begin{cases} \frac{\partial u}{\partial t} = \nabla \cdot (D(S_\rho(\nabla u_\sigma))\nabla u), \ (x,t) \in \Omega \times (0,T], \\ \\ u(x,0) = f, \quad x \in \Omega, \\ \\ \langle D(S_\rho(\nabla u_\sigma))\nabla u, n\rangle = 0, \quad (x,t) \in \partial\Omega \times (0,T]. \end{cases} \tag{1.6.1}$$

where $D(S_\rho(\nabla u_\sigma))$ is symmetric, positive semidefinite matrix, called the diffusion tensor and it is constructed in the following way.

To avoid false detections of edges due to the presence of noise, u is convolve with the Gaussian kernel G_σ and calculated the matrix as follows

$$S_0(\nabla u_\sigma) := \nabla u_\sigma^T \nabla u_\sigma. \tag{1.6.2}$$

This matrix possesses an orthogonal basis composed of eigenvectors v_1, v_2 with $v_1 \parallel \nabla u_\sigma$ and $v_2 \perp \nabla u_\sigma$. The corresponding eigenvalues are equal to $|\nabla u_\sigma|^2$ and 0, respectively, and give contrast in direction of eigenvectors.

In next step, the local information is averaged by convolving S_0 component wise with the Gaussian kernel G_ρ. As a result we obtain symmetric, positive semidefinite matrix

$$S_\rho(\nabla u_\sigma) := G_\rho * S_0(\nabla u_\sigma) := \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}. \tag{1.6.3}$$

The matrix S_ρ is called the structure tensor and possesses orthonormal eigenvectors v_1, v_2 with v_1 parallel to

$$\begin{bmatrix} 2s_{12} \\ \\ s_{22} - s_{11} + \sqrt{(s_{11} - s_{22})^2 + 4s_{12}^2} \end{bmatrix}.$$

The corresponding eigenvalues are given by

$$\mu_1 = \frac{1}{2} \begin{bmatrix} s_{11} + s_{22} + \sqrt{(s_{11} - s_{22})^2 + 4s_{12}^2} \\ \\ \end{bmatrix}$$

and

$$\mu_2 = \frac{1}{2} \begin{bmatrix} s_{11} + s_{22} - \sqrt{(s_{11} - s_{22})^2 + 4s_{12}^2} \\ \\ \end{bmatrix}.$$

This diffusion tensor $D(S_\rho(\nabla u_\sigma))$ should have the same set of eigenvectors as the structure tensor S_ρ, and the choice of corresponding eigenvalues should depend on the desired goal of the filter.

Diffusion tensor. Let us now describe how to define the diffusion tensor $D(S_\rho(\nabla u_\sigma))$ such that the anisotropic filter (1.6.1) could be applied to particular problems.

Since the eigenvectors of the diffusion tensor should reflect the local image structure, one should choose the same orthonormal basis of eigenvectors that one gets from the structure tensor S_ρ. Following Weickert [78], introduce here two possible choices of eigenvalues λ_1 and λ_2 of $D(S_\rho(\nabla u_\sigma))$.

Edge-enhancing anisotropic diffusion. Suppose one wants to smooth an image within some region and at the same time preserve edges, then one should reduce the

diffusivity λ_1 perpendicular to edges more if the contrast given by the greatest eigenvalue μ_1 of the structure tensor S_ρ is large. This behavior may be accomplished by the following choice of eigenvalues,

$$\lambda_1 := \begin{cases} 1, & \text{if } \mu_1 \leq 0 \\ 1 - \exp\left(\frac{-3.315}{\mu_1^4}\right), & \text{if } \mu_1 > 0 \end{cases}$$

$$\lambda_2 := 1.$$

Coherence-enhancing anisotropic diffusion. If one wants to enhance coherent structures, then one should perform smoothening, preferably along the coherence direction v_2 with diffusivity λ_2, which is increasing with respect to the coherence $(\mu_1 - \mu_2)^2$. This may be achieved by the following choice of eigenvalues of the diffusion tensor,

$$\lambda_1 := \alpha$$

$$\lambda_2 := \begin{cases} \alpha, & \text{if } \mu_1 = \mu_2 \\ \alpha + (1-\alpha)\exp\left(\frac{-1}{(\mu_1-\mu_2)^2}\right), & \text{otherwise} \end{cases}$$

where $\alpha \in (0, 1)$ is a small positive parameter which keeps the diffusion tensor $D(S_\rho(\nabla u_\sigma))$ uniformly positive definite.

1.7 Explicit discretization scheme

When applied to discrete signals, the partial differential equation (1.5.1) has to be discretized. In this section, we focus on explicit finite difference schemes. Substituting the spatial partial derivatives in (1.5.1) by finite differences (with the assumption of unit distance between neighboring pixels), and employing explicit discretization in time, an explicit 1-D scheme for nonlinear diffusion can be written in the form

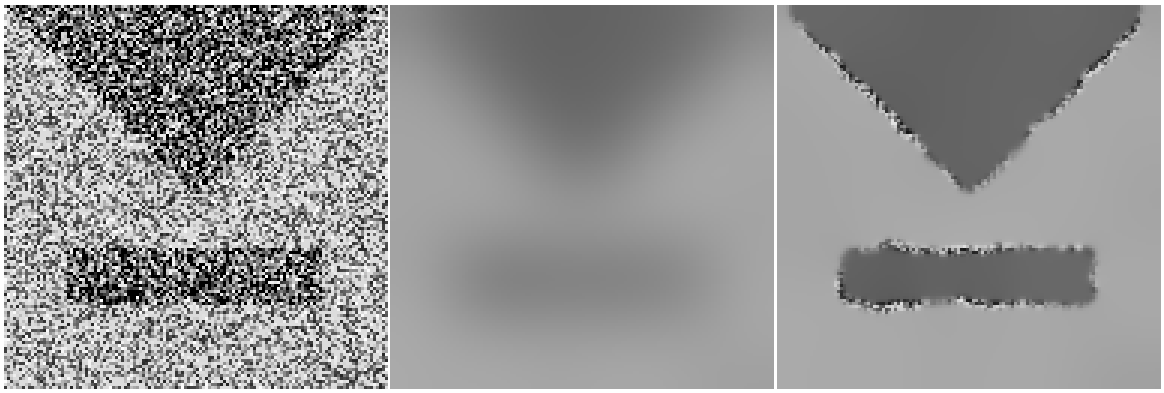
$$\frac{u_i^{k+1} - u_i^k}{\tau} = d(|u_{i+1}^k - u_i^k|)(u_{i+1}^k - u_i^k) - d(|u_i^k - u_{i-1}^k|)(u_i^k - u_{i-1}^k),$$

where τ is the time step size and the upper index k denotes the approximate solution at time $k\tau$. Separating the unknown u_i^{k+1} on one side, we obtain

$$u_i^{k+1} = u_i^k - \tau d(|u_i^k - u_{i+1}^k|)(u_i^k - u_{i+1}^k) + \tau d(|u_{i-1}^k - u_i^k|)(u_{i-1}^k - u_i^k). \qquad (1.7.1)$$

The initial condition reads $u_i^0 = f_i$ for all i.

1.8 Computational result



(a) (b) (c)

Figure 1.8.1: Restoration properties of diffusion filters. (a) left: Noisy image, (b) middle: Image is processed for linear diffusion (c) right : Nonlinear isotropic diffusion

1.9 Correspondence between diffusion and shrinkage functions

Wavelet shrinkage

The discrete wavelet transform represents a one dimensional signal in terms of shifted versions of a dilated lowpass scaling function φ, and shifted and dilated versions of a bandpass wavelet function ψ. In case of orthonormal wavelets, this gives

$$f = \sum_{i\in\mathbb{Z}} \langle f, \varphi_i^n \rangle \varphi_i^n + \sum_{j=-\infty}^{n} \sum_{i\in\mathbb{Z}} \langle f, \psi_i^j \rangle \psi_i^j. \tag{1.9.1}$$

Here we have taken the convention, $\psi_i^j(s) := 2^{-j/2}\psi(2^{-j}s - i)$ and $\langle ., . \rangle$ denotes the inner product in $L^2(\mathbb{R})$. If the measurement f is corrupted by moderate white Gaussian noise, then this noise is contained to a small amount in all wavelet coefficients $\langle f, \psi_i^j \rangle$, while the original signal is in general determined by a few significant wavelet coefficients [43]. Therefore, wavelet shrinkage attempts to eliminate noise from the wavelet coefficients by the following three-step procedure:

1. Analysis: transform the noisy data f to the wavelet coefficients $d_i^j = \langle f, \psi_i^j \rangle$ and scaling function coefficients $c_i^n = \langle f, \varphi_i^n \rangle$.

2. Shrinkage: apply a shrinkage function S_θ with a threshold parameter θ to the wavelet coefficients, i.e., $S_\theta(d_i^j) = S_\theta(\langle f, \psi_i^j \rangle)$.

3. Synthesis: reconstruct the denoised version u of f from the shrunken wavelet coefficients.

$$u := \sum_{i\in\mathbb{Z}} \langle f, \varphi_i^n \rangle \varphi_i^n + \sum_{j=-\infty}^{n} \sum_{i\in\mathbb{Z}} S_\theta(\langle f, \psi_i^j \rangle) \psi_i^j.$$

We restrict our attention to Haar wavelets, well suited for piecewise constant signals with discontinuities. The Haar wavelet and scaling functions are given respectively by

$$\begin{aligned} \psi(x) &= 1_{[0,\frac{1}{2})} - 1_{[\frac{1}{2},1)}, \\ \phi(x) &= 1_{[0,1)}, \end{aligned} \tag{1.9.2}$$

where $1_{[a,b)}$ denotes the characteristic function, equal to 1 on $[a, b)$ and zero everywhere else. Using the so-called "two-scale relation" of the wavelet and its scaling function,

the coefficients c_i^j and d_i^j at higher level j can be computed from the coefficients c_i^{j-1} at lower level $j-1$ and conversely, i.e.,

$$c_i^j = \frac{c_{2i}^{j-1} + c_{2i+1}^{j-1}}{\sqrt{2}}, \quad d_i^j = \frac{c_{2i}^{j-1} - c_{2i+1}^{j-1}}{\sqrt{2}}, \tag{1.9.3}$$

and

$$c_{2i}^{j-1} = \frac{c_i^j + d_i^j}{\sqrt{2}}, \quad c_{2i+1}^{j-1} = \frac{c_i^j - d_i^j}{\sqrt{2}}. \tag{1.9.4}$$

This results in a fast algorithm for the analysis and synthesis steps. Various shrinkage functions leading to qualitatively different denoised functions u were considered in literature, e.g.,

A. Linear shrinkage: $S(x) = \lambda x \qquad (\lambda \in [0,1])$,

B. Soft shrinkage [27]: $S_\theta(x) = \begin{cases} 0 & |x| \le \theta, \\ x - \theta\ \mathrm{sgn}(x) & |x| > \theta. \end{cases}$

C. Garrote shrinkage [30]: $S_\theta(x) = \begin{cases} 0 & |x| \le \theta, \\ x - \frac{\theta^2}{x} & |x| > \theta. \end{cases}$

D. Firm shrinkage [31]: $S_{\theta_1,\theta_2}(x) = \begin{cases} 0 & |x| \le \theta_1, \\ \mathrm{sgn}(x)\frac{\theta_2(|x|-\theta_1)}{\theta_1-\theta_2} & \theta_1 < |x| \le \theta_2, \\ x & \theta_2 < |x|. \end{cases}$

E. Hard shrinkage [43]: $S_\theta(x) = \begin{cases} 0 & |x| \le \theta, \\ x & |x| > \theta. \end{cases}$

1.9.1 Discrete translation invariant scheme

In practice one deals with discrete signals $\mathbf{f} = (f_i)_{i=0}^{N-1}$, where for simplicity N is a power of 2. The Haar wavelet shrinkage starts by setting $c_i^0 = f_i$ and proceeds by analysis, shrinkage and synthesis. Let us just consider a single wavelet decomposition level, i.e., we set $n = 1$. Then, using the convention that $c_i = c_i^1$ and $d_i = d_i^1$, we can drop the superscripts $j = 0$ and $j = 1$. By (1.9.3) and (1.9.4), Haar wavelet shrinkage on one level produces the signal $\mathbf{u}^+ = (u_i^+)_{i=0}^{N-1}$ with coefficients

$$u_{2i}^+ = \frac{c_i + S_\theta(d_i)}{\sqrt{2}} = \frac{f_{2i} + f_{2i+1}}{2} + \frac{1}{\sqrt{2}} S_\theta\left(\frac{f_{2i} - f_{2i+1}}{\sqrt{2}}\right), \tag{1.9.5}$$

$$u^+_{2i+1} = \frac{c_i - S_\theta(d_i)}{\sqrt{2}} = \frac{f_{2i} + f_{2i+1}}{2} - \frac{1}{\sqrt{2}} S_\theta\left(\frac{f_{2i} - f_{2i+1}}{\sqrt{2}}\right). \tag{1.9.6}$$

Note that the single Haar wavelet shrinkage step (1.9.5)-(1.9.6) decouples the input signal into successive pixel pairs: the pixel at position $2i-1$ has no direct connection to its neighbour at position $2i$, and the procedure is not invariant to translation of the input signal. To overcome this problem, Coifman and Donoho [22] introduced the so called cycle spinning: the input signal is shifted, denoised using wavelet shrinkage, shifted back, and the results of all such shifts are averaged. For our single decomposition level, we need only one additional shift to acquire translation invariance. The shifted Haar wavelet shrinkage yields the signal $\mathbf{u}^- = (u^-_i)_{i=0}^{N-1}$ with coefficients,

$$u^-_{2i-1} = \frac{f_{2i-1} + f_{2i}}{2} + \frac{1}{\sqrt{2}} S_\theta\left(\frac{f_{2i-1} - f_{2i}}{\sqrt{2}}\right),$$

$$u^-_{2i} = \frac{f_{2i-1} + f_{2i}}{2} - \frac{1}{\sqrt{2}} S_\theta\left(\frac{f_{2i-1} - f_{2i}}{\sqrt{2}}\right).$$

Averaging the shifted results, one cycle of shift invariant Haar wavelet shrinkage can be summarized into

$$\begin{aligned} u_i &= \frac{u^-_i + u^+_i}{2} \\ &= \frac{f_{i-1} + 2f_i + f_{i+1}}{4} + \frac{1}{2\sqrt{2}} S_\theta\left(\frac{f_i - f_{i+1}}{\sqrt{2}}\right) - \frac{1}{2\sqrt{2}} S_\theta\left(\frac{f_{i-1} - f_i}{\sqrt{2}}\right) \end{aligned} \tag{1.9.7}$$

In order to derive the relation between the explicit diffusion scheme and translation invariant Haar wavelet shrinkage [51], we rewrite the first iteration step in (1.7.1) using the initial condition $u^0_i = f_i$ and the simplified $u^1_i = u_i$ as

$$\begin{aligned} u_i &= \frac{f_{i-1} + 2f_i + f_{i+1}}{4} + \frac{f_i - f_{i+1}}{4} - \frac{f_{i-1} - f_i}{4} \\ &\quad - \tau d(|f_i - f_{i+1}|)(f_i - f_{i+1}) + \tau d(|f_{i-1} - f_i|)(f_{i-1} - f_i) \\ &= \frac{f_{i-1} + 2f_i + f_{i+1}}{4} \\ &\quad + (f_i - f_{i+1})\left(\frac{1}{4} - \tau d(|f_i - f_{i+1}|)\right) \\ &\quad - (f_{i-1} - f_i)\left(\frac{1}{4} - \tau d(|f_{i-1} - f_i|)\right). \end{aligned} \tag{1.9.8}$$

This coincides with (1.9.7) if and only if

$$\frac{1}{2\sqrt{2}} S_\theta\left(\frac{x}{\sqrt{2}}\right) = x\left(\frac{1}{4} - \tau d(|x|)\right). \tag{1.9.9}$$

Equation (1.9.9) relates the shrinkage function S_θ of wavelet denoising to the diffusivity d of nonlinear diffusion. Provided that relation (1.9.9) holds true, a single step of wavelet shrinkage is equivalent to a single step of explicitly discretized nonlinear diffusion. The following two formulae are derived from (1.9.9) and can be used to obtain a shrinkage function S_θ from a diffusivity d, or vice versa.

$$S_\theta(x) = x(1 - 4\tau d(|\sqrt{2}x|)), \tag{1.9.10}$$

$$d(|x|) = \frac{1}{4\tau} - \frac{\sqrt{2}}{4\tau x} S_\theta\left(\frac{x}{\sqrt{2}}\right). \tag{1.9.11}$$

Chapter 2

Viscosity solution of nonlinear diffusion equation

2.1 Introduction

Crandall et al. [24] introduced the notion of viscosity solutions to nonlinear partial differential equations. In this Chapter, we have studied time dependent models for image restoration. These models are constructed by evolving the Euler-Lagrange equations of the optimization problem. We present proof of the existence, uniqueness and stability for the viscosity solution of model (2.2.7). We have used finite difference method to discretize time dependent models.

Image restoration is a fundamental problem in both image processing and computer vision with numerous applications. An image can be interpreted as a real function defined on Ω, a bounded and open domain of $\mathbb{R}^2$ (for simplicity we will assume Ω to be the square domain henceforth). Our problem is to restore an image which is contaminated with noise and blurred in such a way that the process should recover the edges of the images. Formation of a blurred and noisy image is typically modeled as

$$u_0 = H[u] + n, \tag{2.1.1}$$

where u_0 is the observed image, H is the point spread function (PSF) and also known as blur kernel and n is additive noise. For all practical purposes, n is taken as Gaussian white noise, i.e. the values $n_{i,j}$ of n at the pixels (i, j) are independent random variables, each with a Gaussian distribution of zero mean and variance σ^2. The blur operator H

can be described by a Fredholm first kind integral operator

$$H[u(x,y)] = \int_{\Omega} u(s,r)h(x,s,y,r)ds\ dr, \tag{2.1.2}$$

where h is the so-called point spread function (PSF) associated with H.

Among all linear blurring operators, many are shift-invariant and can be expressed in the form of convolution:

$$\begin{aligned} H[u(x,y)] &= \int_{\Omega} u(s,r)h(x-s,y-r)ds\ dr \\ &= (h*u)(x,y). \end{aligned} \tag{2.1.3}$$

With an added noise n, the observed image u_0 is given by,

$$u_0(x,y) = (h*u)(x,y) + n(x,y). \tag{2.1.4}$$

Model (2.1.4) is applicable to a large variety of image degradation processes. Throughout this Chapter, we consider (2.1.4) as the model of degradation where the PSF h is assumed to be known. Typically, h has the following properties, $h(x,y) \geq 0$, $h \to 0$ as $(x^2+y^2) \to +\infty$, and $\int_{\mathbb{R}^2} h(x,y)dx\ dy = 1$. For any $\alpha > 0$, the so-called Gaussian PSF, defined as

$$h_\alpha(x,y) = \frac{1}{2\pi\alpha^2} \exp^{\frac{-(x^2+y^2)}{2\alpha^2}}, \tag{2.1.5}$$

is an important example that will be used in our numerical experiments.

The image restoration problem is ill-posed; either the linear operator H does not admit inverse or it is near singular, yielding highly noise sensitive solutions. A total variation based approach was developed by Rudin, Osher and Fatemi [61] for denoising to overcome the basic limitations of all smooth regularization algorithms. The TV-based technique uses the L^1 norm of the magnitude of a gradient, thus making discontinuous and non smooth solutions possible (if they belong to the space of functions of a bounded total variation). Marquina and Osher [46] have proposed an improved TV model for deblurring and noise removal based on level set ideas. Linear semi-implicit fixed point procedure is another approach which was devised by Vogel and Oman [76]. Chan, Golub and Mulet [15] introduced primal-dual implicit quadratic method to solve

such models. Chang et al. [17] have developed a fixed point method which is an improved Algebraic multigrid (AMG) method, and gives a good initialization to solve the corresponding linear equations. These methods give good results when treating pure denoising problems, but the methods become highly ill-conditioned in the deblurring and denoising case. As t increases, we approach a restored version of image, and the effect of the evolution results into edge enhancement and smoothing. This solution procedure is that of a parabolic equation with time as an evolution parameter and resembles the gradient-projection method of Rosen [59].

Definition 2.1.1. *Consider the general nonlinear equation*

$$F(x, u(x), Du(x), D^2u(x)) = 0, \quad x \in \Omega, \tag{2.1.6}$$

where $F : \Omega \times \mathbb{R} \times \mathbb{R}^N \times \rho^N \to \mathbb{R}$ is a continuous function and Ω is any open set in $\mathbb{R}^N$. Here ρ^N denotes the space of all $N \times N$ real valued symmetric matrices.

We call F degenerate elliptic if the following condition is satisfied:

$$F(x, r, p, X) \leq F(x, r, p, Y) \text{ whenever } X \geq Y,$$

where $X \geq Y$ means that $X - Y$ is nonnegative definite matrix. It is called proper if it is degenerate elliptic and satisfies the monotonicity condition in the r-variable i.e.,

$$F(x, r, p, X) \leq F(x, s, p, Y) \text{ whenever } r \leq s \text{ and } X \geq Y.$$

From now on, we always assume our equations to be proper, unless otherwise specified. To motive the definition of viscosity solutions, we prove the following preposition which is a variant of classical maximum principle.

Proposition 2.1.1. *Let u be a classical solution of the equation (2.1.6) and $\phi : \Omega \to \mathbb{R}$ be any C^2 function. Then if $u - \phi$ has a local maximum (local minimum) at a point $x_0 \in \Omega$, then*

$$F(x_0, u(x_0), D\phi(x_0), D^2\phi(x_0)) \leq (\geq)0. \tag{2.1.7}$$

Proof. The proof is a simple consequence of the maximum principle. If $u - \phi$ has a local maximum at x_0, then $Du(x_0) = D\phi(x_0)$ and $D^2u(x_0) - D^2\phi(x_0) \leq 0$. Now using the degenerate ellipticity, we conclude.

We now give the definition of the viscosity solution.

Definition 2.1.2. *An upper semicontinuous function (resp., lower semicontinuous function) $u : \Omega \to \mathbb{R}$ is called a viscosity subsolution (resp., viscosity supersolution) if*

$$F(x_0, u(x_0), D\phi(x_0), D^2\phi(x_0)) \leq (resp., \geq)0, \tag{2.1.8}$$

whenever x_0 is local maximum (resp., local minimum) of $u - \phi$ for a smooth function $\phi : \Omega \to \mathbb{R}$.
A continuous function u which is both viscosity sub and super-solution is called a viscosity solution.

Remark 2.1.1. *The definition of viscosity solution is a local one. This means that if u is viscosity subsolution in Ω then it is subsolution also in Ω' where $\Omega' \subset \Omega$.*

Remark 2.1.2. *In the definition of viscosity solution, local maximum can be replaced by global maximum and also by strict local or global maximum. Also C^2 functions can be replaced by smooth functions. Also we can assume that the local maximum is zero. Similar remark applies for supersolutions also.*

Now on, we remove the term viscosity and we simply call subsolution or supersolution unless no confusio arises. We also use the following notation throughout: the function ϕ used in the definition of viscosity solutions are called test functions.

2.2 Total variation based restoration algorithms

The total variation (TV) deblurring and denoising models are based on a variational problem with constraints using the TV norm as a nonlinear nondifferentiable functional. The restoration problem as given in [60, 61], can thus be written as

$$\begin{cases} \text{minimize} \int\limits_{\Omega} |\nabla u| \; dx \; dy = \int\limits_{\Omega} \sqrt{u_x^2 + u_y^2} \; dx \; dy, \\ \\ \text{subject to} \;\; ||h * u - u_0||_{L^2}^2 = |\Omega|\sigma^2. \end{cases} \tag{2.2.1}$$

We can then write the equivalent unconstrained problem as:

$$\min_u \int\limits_{\Omega} |\nabla u| \; dx \; dy + \frac{\lambda}{2} \int\limits_{\Omega} (h * u - u_0)^2 \; dx \; dy. \tag{2.2.2}$$

The Euler-Lagrange equation is given by:

$$0 = -\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda h * (h * u - u_0). \tag{2.2.3}$$

Since (2.2.3) is not well defined at points where $\nabla u = 0$, due to the presence of the term $\frac{1}{|\nabla u|}$, it is common to slightly perturb the TV algorithm to become

$$\int_{\Omega} |\nabla u|_{\beta} \; dx \; dy = \int_{\Omega} \sqrt{u_x^2 + u_y^2 + \beta} \; dx \; dy, \tag{2.2.4}$$

where β is a small positive parameter [17].

The time dependent model, see the reference [60], is given by

$$\begin{cases} \frac{\partial u}{\partial t} = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) - \lambda h * (h * u - u_0) \;\; \text{in} \;\; Q_T \equiv I \times \Omega, \\ \\ \partial_v u = 0 \;\; \text{on} \;\; I \times \partial\Omega, \\ \\ u(x, y, 0) = u_0(x, y), \end{cases} \tag{2.2.5}$$

where $\Omega \subset \mathbb{R}^2$ is bounded with Lipschitz continuous boundary $\partial\Omega; I \equiv [0, T]$ is the time interval. The total variation norm does not penalize discontinuities in u, and thus allows us to recover the edges of the original image.

Sometimes the model (2.2.5) converges very slowly to its steady state since the parabolic term is singular for small gradients. At the same time, the Courant-Friedrichs-Lewy (CFL) restriction for keeping stability must be noticed. In order to regularize the parabolic term, Marquina and Osher [46] multiplied the whole Euler-Lagrange equation (2.2.3) by the magnitude of the gradient of the solution,

$$\begin{cases} \frac{\partial u}{\partial t} = |\nabla u| \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) - |\nabla u| \lambda h * (h * u - u_0) \;\; \text{in} \;\; Q_T \equiv I \times \Omega, \\ \\ \partial_v u = 0 \;\; \text{on} \;\; I \times \partial\Omega, \\ \\ u(x, y, 0) = u_0(x, y), \end{cases} \tag{2.2.6}$$

where $\Omega \subset \mathbb{R}^2$ is bounded with Lipschitz continuous boundary $\partial\Omega$; $I \equiv [0,T]$ is the time interval.

To regularize the parabolic term, we multiply the regularizer term by $g(|\nabla G * u|)$ in (2.2.6) and it turns out to be

$$\begin{cases} \frac{\partial u}{\partial t} = g(|\nabla G * u|)|\nabla u|\nabla \cdot \left(\frac{\nabla u}{|\nabla u|}\right) - |\nabla u|\lambda h * (h * u - u_0), \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{in } Q_T \equiv I \times \Omega, \\ \partial_v u = 0 \ \text{ on } \ I \times \partial\Omega, \\ u(x,y,0) = u_0(x,y), \end{cases} \tag{2.2.7}$$

where $\Omega \subset \mathbb{R}^2$ is bounded with Lipschitz continuous boundary $\partial\Omega$; $I \equiv [0,T]$ is time interval.

We have the classical choice $g(s) = \frac{1}{\sqrt{1+(\frac{s}{e})^2}}$ with a constant e or $g(s) = exp(-s^2)$, see the references [18, 56]. According to the previously stated strategy $g(*)$ has to be a nonnegative monotonically decreasing function with $g(0) = 1$. In this way the diffusion process will mainly take place in the interior of regions, and it will not affect the region boundaries where the magnitude of ∇u is large. Here, we take $g(s) = \frac{1}{\sqrt{1+(\frac{s}{e})^2}}$ and $e = 5$, see the reference [82].

2.3 Existence, uniqueness and stability

Motivated by Alvarez et al. [3], we present the existence, uniqueness and stability for the viscosity solution to the model (2.2.7):

$$\begin{cases} \frac{\partial u}{\partial t} = g(\nabla G * u)a_{ij}(\nabla u)u_{x_i x_j} - \lambda|\nabla u|h * (h * u - u_0), \ x \in \mathbb{R}^2, \ t \in \mathbb{R}_+, \\ u(x,0) = u_0(x), \quad x \in \mathbb{R}^2, \end{cases} \tag{2.3.1}$$

where,

$$h(x,y) = \frac{1}{2\pi\alpha^2}\exp^{\frac{-(x^2+y^2)}{2\alpha^2}}(\alpha > 0), \ a_{ij}(p) = \delta_{ij} - \frac{p_i p_j}{|p|^2},$$
$$g \in C^{1,1}(\mathbb{R}^2, \mathbb{R}), \quad g(p) > 0 \text{ for all } p \text{ in } \mathbb{R}^2, \tag{2.3.2}$$

and u_0 is continuous on $\mathbb{R}^2$.

Next, we have to explain the meaning of (2.3.1) which is a second order parabolic equation with possible high degeneracy and three types of nonlinear terms, namely, a quasilinear term $a_{ij}(\nabla u)u_{x_ix_j}$, nonlocal term $g(\nabla G * u)$ and data term $\lambda|\nabla u|h * (h * u - u_0)$.

That is why it is important to work here with viscosity solutions (Crandall et al. [24] and the references [19, 29, 32, 71] for the case when g=1). This is not the place to explain the use of viscosity solutions; let us only point out that classical solutions are automatically viscosity solutions and that general existence, uniqueness, and approximation results are available for viscosity solutions [24].

We first recall the definition of viscosity sub-solution of (2.3.1). A function $u \in C([0,T] \times \mathbb{R}^2)$ for some $T > 0$ is said to be a viscosity sub-solution of (2.3.1), if for all $\phi \in C^2(\mathbb{R}^2 \times \mathbb{R})$, the following condition holds at any point $(x_0, t_0) \in \mathbb{R}^2 \times (0, T]$, for which $(u - \phi)$ attains a local maximum [19].

$$\begin{aligned} &\frac{\partial \phi}{\partial t}(x_0,t_0) - g(\nabla G * u(x_0,t_0))a_{ij}(\nabla\phi(x_0,t_0))\phi_{x_ix_j}(x_0,t_0) \\ &+ \lambda|\nabla\phi(x_0,t_0)|h * (h * u - u_0)(x_0,t_0) \le 0, \quad \text{if } \nabla\phi(x_0,t_0) \neq 0, \end{aligned} \tag{2.3.3}$$

$$\begin{aligned} &\frac{\partial \phi}{\partial t}(x_0,t_0) - g(\nabla G * u(x_0,t_0))\limsup_{p\to 0} a_{ij}(p)\phi_{x_ix_j}(x_0,t_0) \le 0, \\ &\qquad\qquad \text{if } \nabla\phi(x_0,t_0) = 0. \end{aligned} \tag{2.3.4}$$

Notice that the above inequality (in the case when $\nabla\phi(x_0,t_0) \neq 0$) is the inequality expected from the classical maximum principle, the other case being a technical variant.

A viscosity super-solution is similarly defined by substituting "local maximum" for "local minimum", "≤ 0" for "≥ 0", and "limsup" for "liminf" in equations (2.3.3) and (2.3.4), respectively. A viscosity solution is a continuous function which is both a sub-solution and a super-solution. Notice that $\nabla G * u$ is in $C([0,T] \times \mathbb{R}^2)$; hence, its value at (x_0, t_0) is meaningful. We now state and prove the main theorem of this chapter.

Theorem 2.3.1. *The equation (2.3.1) has a unique viscosity solution* $u \in C(\mathbb{R}^2 \times [0,T]) \cap L^\infty(0,T;W^{1,\infty}(\mathbb{R}^2))$ *for any* $T \in [0,\infty)$*, provided that* u_0 *is Lipschitz continuous on* $\mathbb{R}^2$*, and if* $v \in C(\mathbb{R}^2 \times \mathbb{R}_+)$ *is a viscosity solution of (2.3.1) with* u_0 *replaced by a Lipschitz continuous function* v_0*, then for all* $T \in [0,\infty)$*, there exists a constant* $C > 0$*, depending only on* u_0*,* v_0 *and* T*, such that*

$$\sup_{0\le t\le T} ||u(x,t) - v(x,t)||_{L^\infty(\mathbb{R}^2)} \le C||u_0 - v_0||_{L^\infty(\mathbb{R}^2)}. \tag{2.3.5}$$

Moreover, $\inf_{\mathbb{R}^2} u_0 \le u(x,t) \le \sup_{\mathbb{R}^2} u_0$.

Proof. If u is a viscosity solution of equation (2.3.1) then,

$$\inf_{\mathbb{R}^2} u_0 \le u(x,t) \le \sup_{\mathbb{R}^2} u_0, \quad \text{on } \mathbb{R}^2 \times \mathbb{R}_+. \tag{2.3.6}$$

Let $\phi = \sup_{\mathbb{R}^2} u_0 + \delta t (\delta > 0)$ in (2.3.3) and assume that $u - \phi$ attains a local maximum at (x_0,t_0) with $t_0 > 0$, then $\nabla\phi(x_0,t_0) = 0$ and from (2.3.3), $\frac{\partial\phi}{\partial t}(x_0,t_0) \le 0$. This contradicts $\frac{\partial\phi}{\partial t}(x_0,t_0) \equiv \delta > 0$ on $\mathbb{R}^2 \times [0,\infty)$. Therefore, $u - \phi$ must attains its maximum at $t_0 = 0$. So,
$u - \phi \le \sup(u_0 - \sup_{\mathbb{R}^2} u_0)$, then $u \le \sup_{\mathbb{R}^2} u_0 + \delta t$.
Similarly we have
$u \ge \inf_{\mathbb{R}^2} u_0 - \delta t$.
When $\delta \to 0$, we can get (2.3.6).

First, we show that a uniform estimate for the gradient of the viscosity solution of (2.3.1), i.e.,

$$||Du(t,.)||_{L^\infty(\mathbb{R}^2)} \le e^{Ct}||Du_0||_{L^\infty(\mathbb{R}^2)}, \tag{2.3.7}$$

where C depends on u_0, $\sup_{|p|\le R}|\nabla^2 g_\epsilon(p)|$ and $\sup_p |a^\epsilon_{ij}(P)|$ with
$R = ||w||_{L^\infty(\mathbb{R}^2)}||\nabla G||_{L^1(\mathbb{R}^2)}$. Everywhere below, C will denote positive constants depending only on these quantities.
We consider a smooth solution u^ϵ of

$$\begin{cases} \frac{\partial u^\epsilon}{\partial t} = g_\epsilon(\nabla G * u^\epsilon) a^\epsilon_{ij}(\nabla u^\epsilon) u^\epsilon_{x_i x_j} - \lambda b^\epsilon(|\nabla u^\epsilon|) h * (h * u^\epsilon - u^\epsilon_0), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad x \in \mathbb{R}^2,\ t \in \mathbb{R}_+, \\ u^\epsilon(x,0) = u^\epsilon_0(x), \quad x \in \mathbb{R}^2. \end{cases} \tag{2.3.8}$$

We prove a priori estimate on ∇u. This estimate will be formal at that level and will be justified later. In fact, we consider a smooth solution u^ϵ of

$$\begin{cases} \frac{\partial u^\epsilon}{\partial t} = g_\epsilon(\nabla G * w)a_{ij}^\epsilon(\nabla u^\epsilon)u^\epsilon_{x_ix_j} - \lambda b^\epsilon(|\nabla u^\epsilon|)h * (h * u^\epsilon - u_0^\epsilon), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad x \in \mathbb{R}^2, t \in \mathbb{R}_+, \\ u^\epsilon(x,0) = u_0^\epsilon(x), \quad x \in \mathbb{R}^2, \end{cases} \tag{2.3.9}$$

where,

$$\begin{gathered} h(x,y) = \frac{1}{2\pi\alpha^2}\exp^{\frac{-(x^2+y^2)}{2\alpha^2}}, \\ 0 < \epsilon < 1, \\ a_{ij}^\epsilon(p) = (\epsilon+1)\delta_{ij} - \frac{p_ip_j}{|p|^2+\epsilon^2}, \\ b^\epsilon(p) = \sqrt{|p|^2+\epsilon}, \quad g_\epsilon = g + \epsilon, \quad w \in L^\infty(\mathbb{R}^2 \times (0,\infty]), \end{gathered} \tag{2.3.10}$$

$u_0^\epsilon(x) \in C^\infty(\mathbb{R}^2)$ (antireflective) such that $u_0^\epsilon \to u_0$ uniformly and

$$||\nabla u_0^\epsilon||_{L^\infty(\mathbb{R}^2)} \le ||\nabla u_0||_{L^\infty(\mathbb{R}^2)}, ||u_0^\epsilon||_{L^\infty(\mathbb{R}^2)} \le ||u_0||_{L^\infty(\mathbb{R}^2)}.$$

By the theory of quasi-linear uniformly parabolic equations, the problems (2.3.8), (2.3.9) and (2.3.10) admit a smooth solution $u^\epsilon \in C^\infty(\mathbb{R}^2 \times \mathbb{R}_+)$. Since any smooth solution is a viscosity solution, according to (2.3.5), we know that, $|u^\epsilon| \le M$ for $(x,t) \in \mathbb{R}^2 \times [0,\infty)$, where $M > 0$ is a constant depending only on u_0.

Differentiating (2.3.9) with respect to x_k, then multiplying the resulting equation by $2u^\epsilon_{x_k}$ and taking a summation w.r.t. k, we get

$$\begin{aligned} &\frac{\partial|\nabla u^\epsilon|^2}{\partial t} - g_\epsilon(\nabla G * w)a_{ij}^\epsilon(\nabla u^\epsilon)\frac{\partial^2|\nabla u^\epsilon|^2}{\partial x_i\partial x_j} - g_\epsilon(\nabla G * w)\frac{\partial a_{ij}^\epsilon}{\partial l}(\nabla u^\epsilon)u^\epsilon_{x_ix_j}\frac{\partial|\nabla u^\epsilon|^2}{\partial x_l} \\ &\qquad + \lambda\frac{\partial b^\epsilon(\nabla u^\epsilon)}{\partial m}h * (h * u^\epsilon - u_0^\epsilon)\frac{\partial|\nabla u^\epsilon|^2}{\partial x_m} \\ &= -2g_\epsilon(\nabla G * w)a_{ij}^\epsilon(\nabla u^\epsilon)u^\epsilon_{x_kx_i}u^\epsilon_{x_kx_j} + 2\frac{\partial g_\epsilon}{\partial l}(\nabla G * w).(G_{x_lx_k} * w)a_{ij}^\epsilon(\nabla u^\epsilon)u_k^\epsilon u^\epsilon_{x_ix_j} \\ &\qquad - 2\lambda b^\epsilon(\nabla u^\epsilon)h * (h * u^\epsilon_{x_k} - (u_0)^\epsilon_{x_k})u^\epsilon_{x_k}. \end{aligned} \tag{2.3.11}$$

From the definitions of a_{ij}^ϵ, b^ϵ, g^ϵ and h, we have

$$|a_{ij}^\epsilon(\nabla u^\epsilon)u_{x_ix_j}^\epsilon| \leq C(a_{ij}^\epsilon(\nabla u^\epsilon)u_{x_kx_i}^\epsilon u_{x_kx_j}^\epsilon)^{\frac{1}{2}}, \quad \sup_{\mathbb{R}^2} h \leq C,$$

$$\sup_{\mathbb{R}^2} |Db^\epsilon(s)| \leq C, \quad |G_{x_lx_k} * w| \leq C,$$

$$|\frac{\partial g_\epsilon}{\partial l}(\nabla G * w)| \leq C(g_\epsilon(\nabla G * w))^{\frac{1}{2}},$$

where $C > 0$ is a constant depending on $\sup|w|$, g_ϵ and M.

Inserting these estimates into (2.3.11) and using Cauchy's inequality we have

$$\frac{\partial|\nabla u^\epsilon|^2}{\partial t} - g_\epsilon(\nabla G * w)a_{ij}^\epsilon(\nabla u^\epsilon)\frac{\partial^2|\nabla u^\epsilon|^2}{\partial x_i\partial x_j} - g_\epsilon(\nabla G * w)\frac{\partial a_{ij}^\epsilon}{\partial l}(\nabla u^\epsilon)u_{x_ix_j}^\epsilon\frac{\partial|\nabla u^\epsilon|^2}{\partial x_l}$$

$$+\lambda\frac{\partial b^\epsilon(\nabla u^\epsilon)}{\partial m}h * (h * u^\epsilon - u_0^\epsilon)\frac{\partial|\nabla u^\epsilon|^2}{\partial x_m}$$

$$\leq C(|\nabla u^\epsilon|^2 + 1) \text{ in } \mathbb{R}^2 \times \mathbb{R}_+, \tag{2.3.12}$$

where $C > 0$ is a constant depending on $\sup|w|$, g_ϵ and M.

We then deduce easily (2.3.7) by applying the maximum principle [12]. In order to conclude, we only have to approximate (2.3.8) by a (slightly) simpler one of a similar form for which we will be able to produce smooth solutions. Then, we will conclude using the above a priori estimate (which will be valid on the approximated solutions).

By the maximum principle [12], (2.3.12) yields for all $t \in [0, T]$ (for any $T < \infty$),

$$||\nabla u^\epsilon(., t)||_{L^\infty(\mathbb{R}^2)} \leq e^{Ct}||\nabla(u_0)^\epsilon||_{L^\infty(\mathbb{R}^2)} \leq e^{Ct}||\nabla(u_0)||_{L^\infty(\mathbb{R}^2)} \leq C_T,$$

where C_T denotes various constants independent of $\epsilon,\ t,\ x,\ y$. This implies that

$$|u^\epsilon(x, t) - u^\epsilon(y, t)| \leq C_T|x - y|, \quad \text{for } \forall\ x, y \in \mathbb{R}^2 \text{ and } \forall\ t \in [0, T].$$

By the same argument, we have

$$|u^\epsilon(x, s) - u^\epsilon(x, t)| \leq C_T|s - t|^{\frac{1}{2}}, \quad \text{for } \forall\ x \in \mathbb{R}^2 \text{ and } \forall\ s, t \in [0, T].$$

Then by Ascoli-Arzela theorem, there exists a subsequence u^{ϵ_k} of u^ϵ, and a function

$u \in C(\mathbb{R}^2 \times [0,T]) \cap L^\infty(0,T;W^{1,\infty}(\mathbb{R}^2))$ such that as $\epsilon_k \to 0$,

$$u^{\epsilon_k} \to u \ \text{ locally uniformly in } \mathbb{R}^2 \times \mathbb{R}_+. \tag{2.3.13}$$

So it is easily to get (2.3.7).

Second, we shall show a existence of viscosity solution.
We will show that u obtained in (2.3.13) is a viscosity solution of (2.3.1) in the sense of (2.3.3) and (2.3.4). Let $\phi \in C^2(\mathbb{R}^2 \times \mathbb{R}_+)$. Firstly we assume $u - \phi$ has a strict local maximum at a point $(x_0, t_0) \in \mathbb{R}^2 \times \mathbb{R}_+$. When $u^{\epsilon_k} \to u$ uniformly near (x_0, t_0), $u^{\epsilon_k} - \phi$ has a local maximum at a point (x_k, t_k) with

$$(x_k, t_k) \to (x_0, t_0), \quad k \to \infty \tag{2.3.14}$$

and

$$\nabla u^{\epsilon_k} = \nabla \phi, \qquad \frac{\partial u^{\epsilon_k}}{\partial t} = \frac{\partial \phi}{\partial t}, \quad a_{ij}^{\epsilon_k}(\nabla u^{\epsilon_k}) u_{x_i x_j}^{\epsilon_k} \le a_{ij}^{\epsilon_k}(\nabla \phi)\phi_{x_i x_j}.$$

Therefore, (2.3.8) implies that at (x_k, t_k),

$$\frac{\partial \phi}{\partial t} - g_{\epsilon_k}(\nabla G * u^{\epsilon_k}) a_{ij}^{\epsilon_k}(\nabla \phi)\phi_{x_i x_j} + b^{\epsilon_k}(\nabla \phi)(u^{\epsilon_k} - (u_0^{\epsilon_k})) \le 0. \tag{2.3.15}$$

(1) If $\nabla \phi(x_0, t_0) \neq 0$, according to (2.3.14), $\nabla \phi(x_k, t_k) \neq 0$ for sufficiently large k. So applying limits to (2.3.15), we can get

$$\frac{\partial \phi}{\partial t} - g(\nabla G * u) a_{ij}(\nabla \phi)\phi_{x_i x_j} + b(\nabla \phi)(u - (u_0)) \le 0, \quad \text{at } (x_0, t_0), \tag{2.3.16}$$

which is the same as (2.3.1).

(2) If $\nabla \phi(x_0, t_0) = 0$, according to (2.3.14), $\nabla \phi(x_k, t_k) \to 0, \epsilon \to 0$ as $k \to 0$, hence, $b^{\epsilon_k}(\nabla \phi(x_k, t_k)) \to 0$. Recalling the definition of $a_{ij}^{\epsilon_k}, b^\epsilon$, and g_ϵ, (2.3.15) can be written as

$$\begin{aligned} \frac{\partial \phi}{\partial t} - (g(\nabla G * u) + \epsilon_k) \left((\epsilon_k + 1)\delta_{ij} - \frac{(\nabla \phi)_i (\nabla \phi)_j}{|\nabla \phi|^2 + \epsilon^2} \right) \phi_{x_i x_j} \\ + b^{\epsilon_k}(\nabla \phi)(u^{\epsilon_k} - (u_0^{\epsilon_k})) \le 0, \quad \text{at } (x_k, t_k). \end{aligned} \tag{2.3.17}$$

Applying limits to (2.3.17), we get

$$\frac{\partial \phi}{\partial t} - g(\nabla G * u) \left(\delta_{ij} - \frac{(\nabla \phi)_i (\nabla \phi)_j}{|\nabla \phi|^2 + \epsilon^2} \right) \phi_{x_i x_j} \le 0, \quad \text{at } (x_0, t_0),$$

which is the same as (2.3.4).
Secondly if $u-\phi$ has a local maximum, but not necessary a strict local maximum at (x_0,t_0), we only need to use $\hat{\phi}(x,t)=\phi(x,t)+|x-x_0|^4+(t-t_0)^4$ to replace $\phi(x,t)$. The process of proof is the same.

Therefore, u is a sub-solution of (2.3.1). Similarly, we can show that u is a super-solution. Hence, u is a viscosity solution of (2.3.1).

Third, we shall show stability.
Let u_0 and v_0 be Lipschitz continuous initial data and u and v be the corresponding viscosity solution of (2.3.1) respectively. Let

$$\omega(x,y,t)=u(x,t)-v(y,t)-(4\delta)^{-1}|x-y|^4-\lambda t, \quad t\in[0,T],\ x,y\in\mathbb{R}^2,$$

where $\delta>0$ and $\lambda>0$ are constants to be determined later. We assume that $\omega(x,y,t)$ attains its maximum at a point (x_0,y_0,t_0) with $t_0>0$. By the arguments given in [24], there exist X and Y, $(n\times n)$ $(n\in Z^+)$ symmetric matrices and $a,b\in\mathbb{R}$, such that

$$a-b=\lambda, \quad \begin{pmatrix} X & 0\\ 0 & -Y\end{pmatrix}\le\begin{pmatrix} A+\mu A^2 & -A-\mu A^2\\ -A-\mu A^2 & A+\mu A^2\end{pmatrix}, \tag{2.3.18}$$

for each $\mu>0$, and

$$\begin{aligned} a-g((\nabla G*u)(x_0,t_0))a_{ij}(\delta^{-1}|x_0-y_0|^2(x_0-y_0))X_{ij}\\ +\delta^{-1}|x_0-y_0|^3(u(x_0,t_0)-u_0(x_0))\le 0,\end{aligned} \tag{2.3.19}$$

$$\begin{aligned} b-g((\nabla G*v)(y_0,t_0))a_{ij}(\delta^{-1}|x_0-y_0|^2(x_0-y_0))Y_{ij}\\ +\delta^{-1}|x_0-y_0|^3(v(y_0,t_0)-v_0(y_0))\ge 0,\end{aligned} \tag{2.3.20}$$

where,

$$A=(A_{ij})_{n\times n}=\delta^{-1}|x_0-y_0|^2I_n+2\delta^{-1}(x_0-y_0)\otimes(x_0-y_0),$$

so that

$$A^2=(A_{ij})_{n\times n}=\delta^{-2}|x_0-y_0|^4I_n+8\delta^{-2}|x_0-y_0|^2(x_0-y_0)\otimes(x_0-y_0),$$

where I_n is element matrix.

Since (x_0, y_0, t_0) is the maximum point of $\omega(x, y, t)$, we can get

$$u(x_0, t_0) - v(y_0, t_0) - (4\delta)^{-1}|x_0 - y_0|^4 - \lambda t_0 \geq u(y_0, t_0) - v(y_0, t_0) - \lambda t_0.$$

This leads to

$$|x_0 - y_0| \leq (4\delta L)^{\frac{1}{3}}, \tag{2.3.21}$$

where L is a Lipschitz constant for u in $\mathbb{R}^2 \times [0, T]$.

If $x_0 = y_0$, then A=0 so that by (2.3.18), $X \leq 0,\ Y \geq 0$. We then get $a \leq 0,\ b \geq 0$, which contradicts $a - b = \lambda > 0$. Therefore $x_0 \neq y_0$.
We now let $\mu = \delta|x_0 - y_0|^2$, and deduce

$$\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq 2\delta^{-1} \begin{pmatrix} B & -B \\ -B & B \end{pmatrix}, \tag{2.3.22}$$

where

$$B = |x_0 - y_0|^2 I_n + 5(x_0 - y_0) \otimes (x_0 - y_0),$$

where I_n is element matrix as above.

We then set

$$g_1 = g((\nabla G * u)(x_0, t_0)), \quad g_2 = g((\nabla G * v)(y_0, t_0)),$$

$$a = a_{ij}(\delta^{-1}|x_0 - y_0|^2(x_0 - y_0))_{1 \leq i,j \leq n},$$

and consider the matrix

$$\tau = \begin{pmatrix} g_1 a & (g_1 g_2)^{\frac{1}{2}} a \\ (g_1 g_2)^{\frac{1}{2}} a & g_2 a \end{pmatrix}. \tag{2.3.23}$$

Obviously, τ is a nonnegative symmetric matrix so that multiplying (2.3.22) to the left by τ and taking the trace, we get

$$\begin{aligned} g_1 a_{ij} X_{ij} - g_2 a_{ij} Y_{ij} &\leq 2\delta^{-1}(g_1^{\frac{1}{2}} - (g_2^{\frac{1}{2}})^2 \mathrm{trace}(aB) \\ &\leq C_0 \delta^{-1}(g_1^{\frac{1}{2}} - (g_2^{\frac{1}{2}})^2 |x_0 - y_0|^2, \end{aligned} \tag{2.3.24}$$

for some C_0 which depends only on $(a_{ij}(p))_{1 \leq i,j \leq n}$.

Combining (2.3.18)-(2.3.20) and (2.3.24), we obtain

$$\lambda = a - b \leq C_0\delta^{-1}(g_1^{\frac{1}{2}} - (g_2^{\frac{1}{2}})^2|x_0 - y_0|^2 + \delta^{-1}|x_0 - y_0|^3[(-u(x_0,t_0) + v(y_0,t_0)) + ((u_0)(x_0) - (v_0)(y_0))]$$

$$\leq C_0\delta^{-1}(g_1^{\frac{1}{2}} - (g_2^{\frac{1}{2}})^2|x_0 - y_0|^2 + \delta^{-1}|x_0 - y_0|^3[|u(x_0,t_0) - v(y_0,t_0)| + |(u_0)(x_0) - (v_0)(y_0)|]. \tag{2.3.25}$$

We now estimate $(g_1^{\frac{1}{2}} - g_2^{\frac{1}{2}})$. First of all, we observe that (2.3.2) yields that $g^{\frac{1}{2}}$ is Lipschitz on bounded sets, therefore

$$(g_1^{\frac{1}{2}} - g_2^{\frac{1}{2}}) \leq C_1|(\nabla G * u)(x_0,t_0) - (\nabla G * v)(y_0,t_0)|,$$

for some C_1, depending only on g and on $\sup|u|$, $\sup|v|$.

This allows us to deduce from (2.3.25) that

$$\lambda \leq C\delta^{-1}\left\{\left(\left(\sup_{\mathbb{R}^2\times[0,T]}|u-v|\right)^2|x_0-y_0|^2\right) + \left(\left(\sup_{\mathbb{R}^2\times[0,T]}|u-v|\right)|x_0-y_0|^3 + |x_0-y_0|^4\right)\right\}, \tag{2.3.26}$$

where $C = 2C_0C_1^2$.

From (2.3.21), therefore (2.3.26) finally yield

$$\lambda \leq M\left\{\left(\sup_{\mathbb{R}^2\times[0,T]}|u-v|\right)^2\delta^{\frac{-1}{3}} + \sup_{\mathbb{R}^2\times[0,T]}|u-v| + \delta^{\frac{1}{3}}\right\}, \tag{2.3.27}$$

where $M = \max((4L)^{\frac{2}{3}}, 4L, (4L)^{\frac{4}{3}})C$.

We now set

$$\delta^{\frac{1}{3}} = \gamma \sup_{\mathbb{R}^2\times[0,T]}|u-v|, \quad \gamma > 0 \tag{2.3.28}$$

and from (2.3.27) we obtain

$$\lambda \leq (1 + \gamma + \frac{1}{\gamma})M \sup_{\mathbb{R}^2\times[0,T]}|u-v|. \tag{2.3.29}$$

If we let

$$\lambda \leq (2+\gamma+\frac{1}{\gamma})M \sup_{\mathbb{R}^2\times[0,T]}|u-v|. \tag{2.3.30}$$

This would lead to contradiction with (2.3.29). Therefore $\omega(x,y,t)$ attains its maximum at $t=0$ for the values of λ in (2.3.30). Hence,
$u(x,t)-v(y,t)-(4\delta)^{-1}|x-y|^4-\lambda t$

$$\begin{aligned}
&\leq \sup_{\mathbb{R}^2\times[0,T]}(u(x,0)-v(y,0)-(4\delta)^{-1}|x-y|^4)\\
&\leq \sup_{\mathbb{R}^2\times[0,T]}(u_0(x)-u_0(y)+u_0(y)-v_0(y)-(4\delta)^{-1}|x-y|^4)\\
&\leq \sup_{\mathbb{R}^2}|u_0-v_0|+\sup_{|x-y|\geq 0}(u_0(x)-u_0(y)-(4\delta)^{-1}|x-y|^4)\\
&\leq \sup_{\mathbb{R}^2}|u_0-v_0|+\sup_{|x-y|\geq 0}(L|x-y|-(4\delta)^{-1}|x-y|^4)\\
&\leq \sup_{\mathbb{R}^2}|u_0-v_0|+(L(\delta L)^{\frac{1}{3}}-(4\delta)^{-1}(\delta L)^{\frac{4}{3}})
\end{aligned}$$

$$\leq \sup_{\mathbb{R}^2}|u_0-v_0|+\frac{3}{4}L^{\frac{4}{3}}\delta^{\frac{1}{3}}=\sup_{\mathbb{R}^2}|u_0-v_0|+\frac{3}{4}L^{\frac{4}{3}}\ \gamma \sup_{\mathbb{R}^2\times[0,T]}|u-v|, \tag{2.3.31}$$

because that $\sup_{r\geq 0}(Lr-(4\delta)^{-1}r^4)$ is achieved at $r=(\delta L)^{\frac{1}{3}}$, and letting x=y in (2.3.31), from (2.3.30), we can get

$$\sup_{\mathbb{R}^2\times[0,T]}|u-v|\leq \sup_{\mathbb{R}^2}|u_0-v_0|+\frac{3}{4}L^{\frac{4}{3}}\gamma \sup_{\mathbb{R}^2\times[0,T]}|u-v|+M(2+\gamma+\gamma^{-1})T\sup_{\mathbb{R}^2\times[0,T]}|u-v|. \tag{2.3.32}$$

Exchanging the role of u and v, and chooching $\gamma=L^{\frac{-4}{3}}$, we deduce

$$\sup_{\mathbb{R}^2\times[0,T]}|u-v|\leq 4\sup_{\mathbb{R}^2}|u_0-v_0|+KT\sup_{\mathbb{R}^2\times[0,T]}|u-v|, \tag{2.3.33}$$

where $K=4M(2+\gamma+\gamma^{-1})$. In order to conclude, we choose $T=t=\frac{1}{2K}$ and we find

$$\sup_{\mathbb{R}^2\times[0,T]}|u-v|\leq 8\sup_{\mathbb{R}^2}|u_0-v_0|. \tag{2.3.34}$$

For large $t=NT(N\geq 1)$, by iteration, we easily obtain

$$\sup_{\mathbb{R}^2\times[0,T]}|u-v|\leq 8^N\sup_{\mathbb{R}^2}|u_0-v_0|. \tag{2.3.35}$$

This proves the stability for u.

At last, we shall show uniqueness.

Suppose u and $\overline{u}$ are two solutions of (2.3.1). According to (2.3.35), we can easily get

$$\sup_{\mathbb{R}^2\times[0,T]} |u-\overline{u}| \leq 8^N \sup_{\mathbb{R}^2} |u_0-u_0| = 0. \tag{2.3.36}$$

So $u=\overline{u}$. This proves the uniqueness of the problem (2.3.1).

2.4 The discrete scheme

Let u_{ij}^n be the approximation to the value $u(x_i, y_j, t_n)$, where

$$x_i = i\Delta x,\ y_j = j\Delta x,\ i,j = 1,2,.....,N,$$

$$N\Delta x = 1, \quad t_n = n\Delta t, \quad n \geq 1,$$

where $\Delta x, \Delta y$ and Δt are the spatial step sizes and the time step size respectively.

The explicit partial derivatives of model (2.2.6) can be expressed as:

$$u_t = \frac{u_{xx}(u_y^2+\beta) - 2u_{xy}u_xu_y + u_{yy}(u_x^2+\beta)}{(u_x^2+u_y^2+\beta)} - \sqrt{u_x^2+u_y^2+\beta}\ \lambda h*(h*u-u_0). \tag{2.4.1}$$

The explicit partial derivatives of model (2.2.7) can be expressed as:

$$\begin{aligned} u_t = g(|\nabla G*u|)\ &\left(\frac{u_{xx}(u_y^2+\beta) - 2u_{xy}u_xu_y + u_{yy}(u_x^2+\beta)}{(u_x^2+u_y^2+\beta)}\right) \\ &- \sqrt{u_x^2+u_y^2+\beta}\ \lambda\ h*(h*u-u_0), \end{aligned} \tag{2.4.2}$$

where $g(|\nabla G*u|) = \frac{e}{(e^2+|\nabla G*u|^2)^{\frac{1}{2}}}$.

We define the derivative terms as,

$$u_{ij}^x = \frac{u_{i+1,j}^n - u_{i-1,j}^n}{2\Delta x};\ u_{ij}^y = \frac{u_{i,j+1}^n - u_{i,j-1}^n}{2\Delta x};\ u_{ij}^{xx} = \frac{u_{i+1,j}^n - 2u_{i,j}^n + u_{i-1,j}^n}{\Delta x^2},$$

$$u_{ij}^{yy} = \frac{u_{i,j+1}^n - 2u_{i,j}^n + u_{i,j-1}^n}{\Delta x^2};\ u_{ij}^{xy} = \frac{u_{i+1,j+1}^n - u_{i-1,j+1}^n - u_{i+1,j-1}^n + u_{i-1,j-1}^n}{4\Delta x\Delta x};\ u_{ij}^t = \frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t}.$$

We let

$$r_{ij}^n = u_{ij}^{xx}((u_{ij}^y)^2+\beta) - 2u_{ij}^{xy}u_{ij}^xu_{ij}^y + u_{ij}^{yy}((u_{ij}^x)^2+\beta);$$

$$w_{ij}^n = \sqrt{(u_{ij}^x)^2 + (u_{ij}^y)^2 + \beta}; \qquad l_{ij}^n = \frac{e}{(e^2 + |\nabla G * u_{ij}^n|^2)^{\frac{1}{2}}}.$$

Then (2.4.1) and (2.4.2) read, respectively as follows:

$$u_{ij}^t = \frac{r_{ij}^n}{(w_{ij}^n)^2} - w_{ij}^n \ \lambda \ h * (h * u_{ij}^n - u_{ij}^0). \tag{2.4.3}$$

$$u_{ij}^t = l_{ij}^n \ \left(\frac{r_{ij}^n}{(w_{ij}^n)^2}\right) - w_{ij}^n \ \lambda \ h * (h * u_{ij}^n - u_{ij}^0), \tag{2.4.4}$$

with boundary conditions

$$u_{i,N}^n = u_{i,N-1}^n, \ u_{N,j}^n = u_{N-1,j}^n, \ u_{1,j}^n = u_{2,j}^n, \ u_{i,1}^n = u_{i,2}^n. \tag{2.4.5}$$

The explicit method is stable and convergent for $\frac{\Delta t}{\Delta x^2} \leq 0.5$, see [41].

For a blurred image, we need to make certain assumptions on the unknown boundary data. Here we use reflective BCs [52], which means the data outside the domain of consideration are taken as a reflection of the data inside. However, the reflection guarantees the continuity, but generally fails to guarantee the continuity of the normal derivative except in the non generic case, where the normal derivative at the boundary is zero. Therefore if the image is smooth, then the reflective BCs save the continuity but introduce an artificial discontinuity of the normal derivative.

2.5 Numerical experiments

We have used three gray scale images, Lena (256×256) and Boat (256×256) as shown in Figure 2.5.1 for our deblurring and denoising experiments. The pixel values of all images lie in interval [0, 255], and the Gaussian white noise is added by the normal imnoise function imnoise (I,'Gaussian', M, σ^2), i.e., the mean M and variance σ^2 in Matlab. We first scale the intensities of the images into the range between zero and one before we begin our experiments.

Blurred signal to noise ratio (BSNR) is used to measure the ratio of the level of blur kernel and the level of noise,

$$\text{BSNR} = 10 \ \log_{10} \left(\frac{\text{blurred signal variance}}{\text{noise variance}}\right) dB. \tag{2.5.1}$$

Improvement in the signal quality (ISNR) is used to measure the goodness of restored image:

$$\text{ISNR} = 10 \log_{10} \left(\frac{\sum_{i,j}^{n} [u_{ij} - (u_0)_{ij}]^2}{\sum_{i,j}^{n} [u_{i,j} - (u_{new})_{i,j}]^2} \right) dB, \tag{2.5.2}$$

where u_{new} is the restored image. That is, the value of ISNR is larger, the restored image is better.

Fig. 2.5.2(a)-(b) and Fig. 2.5.3(a)-(b) represent the blurred and noisy images with the size of Gaussian blur kernel $h_\alpha = 5$, the blurring parameter $\alpha = 2$ and different levels Gaussian white noise $\sigma^2 = 0.003$, $\sigma^2 = 0.005$ respectively. We have taken Lagrange multiplier $\lambda = 0.85$ as was used in references [15, 17]. We can choose $\beta = 10^{-32}$ [17], the smallest positive machine number. We have used in our all experiment $\frac{\Delta t}{\Delta x^2} = 0.1$, see the reference [46].

The BSNR of Figures 2.5.2(a) and 2.5.3(a) are ≈ 11.59 and 10.43 respectively. The BSNR of Figures 2.5.2(b) and 2.5.3(b) are ≈ 9.37 and 8.21 respectively.

The values of ISNR using our model (2.2.7) is larger than that of using model (2.2.6) at similar iteration numbers. Thus it can be concluded that model (2.2.7) get better restored images than that of model (2.2.6) for a degraded image.

(a)

(b)

Figure 2.5.1: Original Test Images used for different experiments (a) Lena: 256×256, (b) Boat: 256×256.

(a)

(b)

(c)

(d)

(e)

(f)

Figure 2.5.2: (top row) Blurred and noisy Lena images with blurring parameter $\alpha = 2$ and different levels of Gaussian noise (a)-(b) $\sigma^2 = 0.003,\ 0.005$, respectively; (second row) (c)-(d) corresponding reconstructed images by model (2.2.6); (e)-(f) by model (2.2.7).

(a)

(b)

(c)

(d)

(e)

(f)

Figure 2.5.3: (top row) Blurred and noisy Boat images with blurring parameter $\alpha = 2$ and different levels of Gaussian noise (a)-(b) $\sigma^2 = 0.003,\ 0.005$, respectively; (second row) (c)-(d) corresponding reconstructed images by model (2.2.6); (e)-(f) by model (2.2.7).

Table 2.5.1: Results obtained by using models (2.2.6) and (2.2.7) applied to the images in Figures 2.5.2(a) and 2.5.3(a) with the size of Gaussian blur operator $h_\alpha = 5$, blurring parameter $\alpha = 2$ and Gaussian noise $\sigma^2 = 0.003$.

Images	ISNR (2.2.6)	Images	ISNR (2.2.7)
Fig. 2.5.2(c)	2.1365	Fig. 2.5.2(e)	2.4667
Fig. 2.5.3(c)	1.9079	Fig. 2.5.3(e)	2.2648
No. of iterations	20	No. of iterations	20

Table 2.5.2: Results obtained by using models (2.2.6) and (2.2.7) applied to the images in Figures 2.5.2(b) and 2.5.3(b) with the size of Gaussian blur operator $h_\alpha = 5$, blurring parameter $\alpha = 2$ and Gaussian noise $\sigma^2 = 0.005$.

Images	ISNR (2.2.6)	Images	ISNR (2.2.7)
Fig. 2.5.2(d)	3.1730	Fig. 2.5.2(f)	3.2862
Fig. 2.5.3(d)	2.9337	Fig. 2.5.3(f)	3.0843
No. of iterations	20	No. of iterations	20

2.6 Conclusion

In this Chapter, we have proposed new time dependent models to solve image restoration problem. The denoising technique is used to every solution image before solving the nonlinear diffusion problem. We have proved the existence, uniqueness and stability of our new model. Nonlinear explicit schemes are used to discretize models (2.2.6) and (2.2.7). Our model (2.2.7) give better restoration results in comparison with other existing model (2.2.6).

Chapter 3

Weak solution of nonlinear diffusion equation

3.1 Introduction

The aim of this Chapter is to study a nonlinear diffusion models for image denoising. We present proof of the existence and uniqueness theorem of model (3.2.1). We derive some properties of weak solution. We have tested our algorithm on various types of images. To quantify the results, we have used peak signal to noise ratio (PSNR) as metric.

Image denoising is a fundamental problem in both image processing and computer vision with numerous applications. The total variation models [15, 61, 76] and anisotropic diffusion models [13, 56, 78, 83, 86] have been studied as a useful tool to the problem of image denoising and image reconstruction. These partial differential equations based image enhancement techniques have been able to achieve a good edge preservation.

Definition 3.1.1. *Let $m > 0$ be an integer and let $1 \leq p \leq \infty$. The Sobolev space $W^{m,p}(\Omega)$ is defined by*

$$W^{m,p}(\Omega) = \{u \in L^p(\Omega) | D^\alpha u \in L^p(\Omega) \quad \textit{for all } |\alpha| \leq m\}, \tag{3.1.1}$$

where Ω is an open set in $\mathbb{R}^n$.

In other words, $W^{m,p}(\Omega)$ is the collection of all functions in $L^p(\Omega)$ such that all distribution derivatives upto order m are also in $L^p(\Omega)$. Clearly $W^{m,p}(\Omega)$ is a vector space.

In all that follows we will consider functions with values in $\mathbb{R}$ and the corresponding function spaces as vector spaces over $\mathbb{R}$. We provide it with the norm:

$$||u||_{m,p,\Omega} = \sum_{|\alpha|\leq m} ||D^\alpha u||_{L^p(\Omega)} \tag{3.1.2}$$

or, equivalently, for $1 < p < \infty$,

$$||u||_{m,p,\Omega} = \left(\sum_{|\alpha|\leq m} \int_\Omega |D^\alpha u|^p \right)^{\frac{1}{p}} = \left(\sum_{|\alpha|\leq m} ||D^\alpha u||^p_{L^p(\Omega)} \right)^{\frac{1}{p}}. \tag{3.1.3}$$

Remark 3.1.1. *The case p=2 will play a special role in the sequel. These spaces will be denoted by $H^m(\Omega)$. Thus*

$$H^m(\Omega) = W^{m,2}(\Omega) \tag{3.1.4}$$

and for $u \in H^m(\Omega)$, we denote its norm by $||u||_{m,\Omega}$. i.e.

$$||u||_{m,\Omega} = ||u||_{m,2,\Omega}. \tag{3.1.5}$$

Remark 3.1.2. *We will also often use the semi-norms which consist of the L^p-norms of the highest order derivatives. We denote these by $|\cdot|_{m,p,\Omega}$. Thus for $u \in W^{m,p}(\Omega)$.*

$$|u|_{m,p,\Omega} = \sum_{|\alpha|=m} ||D^\alpha u||_{L^p(\Omega)}, \tag{3.1.6}$$

with the obvious modification (when $1 < p < \infty$) if we use equation (3.1.3) to define the norm. Consistent to remark (3.1.1) if $p = 2$ we only write $|\cdot|_{m,\Omega}$ instead of $|\cdot|_{m,2,\Omega}$.

Remark 3.1.3. *We can naturally consider the space $L^p(\Omega)$ as a special case of the Sobolev class, viz. when $m = 0$. i.e. we do not bother about derivatives. In particular we denote the L^p-norm of a function by $|\cdot|_{0,p,\Omega}$ (since in this case the semi-norm and norm are the same). Again the $L^2(\Omega)$-norm will be denoted by $|\cdot|_{0,\Omega}$.*

3.2 Nonlinear diffusion model using total variation and Perona-Malik diffusivities

An image can be interpreted as a real function defined on Ω, a bounded and open domain of $\mathbb{R}^2$ (for simplicity we will assume Ω to be the square domain henceforth). Formation of a noisy image is typically modeled as

$$u_0(x) = u(x) + n(x),$$

where $u(x)$ denote the desired clean image, $u_0(x)$ denote the pixel values of a noisy image for $x \in \Omega$ and $n(x)$ is additive white noise assumed to be close to Gaussian. The values $n(i,j)$ of n at the pixels (i,j) are independent random variables, each with a Gaussian distribution of zero mean and variance σ^2.

We propose the following second order - version of the nonlinear diffusion model which is a synthesis of ideas from Catté et al. [13]. Our model is given by:

$$\begin{cases} \frac{\partial u}{\partial t} = \nabla\cdot(g_1(|\nabla G_\sigma * u|)\nabla u) + \nabla\cdot(g_2(|\nabla G_\sigma * u|)\nabla u) - \lambda(u - u_0) \text{ on } \Omega \times (0,T), \\ u(x,0) = u_0(x) \quad \text{in } \Omega, \\ \frac{\partial u}{\partial \vec{n}} = 0 \quad \text{on } \partial\Omega \times (0,T), \end{cases} \tag{3.2.1}$$

where g_1 and g_2 are decreasing function tending to zero at infinity with $g_i(0) = M_i > 0$ and $t \to g_i(\sqrt{t})$ is smooth, and $G_\sigma(x)$ is the Gaussian kernel, namely,

$$G_\sigma(x) = \frac{1}{2\pi\sigma^2} e^{\frac{-|x|^2}{2\sigma^2}}. \tag{3.2.2}$$

3.3 Existence and uniqueness of weak solutions

In this Section, we establish the existence and uniqueness of weak solutions of the proposed model following the arguments in [13, 28]. The standard notations are used throughout.

Let Ω denote the square $(0,1)\times(0,1)$ of $\mathbb{R}^2$ and $H^k(\Omega)$, k a positive integer, the set of all functions $u(x)$ defined on Ω such that u and its distributional derivatives $\frac{\partial^m u}{\partial x^m}$ of order $|m|=\sum_{j=1}^{N} m_j \leq k$ for all belong to $L^2(\Omega)$. $H^k(\Omega)$ is a Hilbert space with the norm

$$||u||_{H^k(\Omega)} = \left(\sum_{|m|\leq k} \int_\Omega \left| \frac{\partial^m u}{\partial x^m} \right|^2 dx \right)^{\frac{1}{2}}. \tag{3.3.1}$$

The space $L^\infty(0,T;H^1(\Omega))$ consists of all functions u such that, for almost every t in (0,T), u belongs to $H^1(\Omega)$. $L^\infty(0,T;H^1(\Omega))$ is a normed space with the norm

$$||u||_{L^\infty(0,T;H^1(\Omega))} = \operatorname*{ess\,sup}_{0\leq t\leq T} ||u(\cdot,t)||_{H^1(\Omega)}. \tag{3.3.2}$$

Theorem 3.3.1. *Let $u_0 \in H^1(\Omega)$ and $||u_0||_{H^1(\Omega)}$ is appropriately small. Then we have a unique weak solution $u(x,t)$ such that $u \in C([0,T];L^2(\Omega)) \cap L^\infty(0,T;H^1(\Omega))$, and verifying*

$$\begin{cases} \frac{\partial u}{\partial t} - \nabla\cdot(g_1(|\nabla G_\sigma * u|)\nabla u) - \nabla\cdot(g_2(|\nabla G_\sigma * u|)\nabla u) = 0 \ \text{in}\ \Omega\times(0,T), \\ u(x,0) = u_0(x) \quad \text{in} \quad \Omega, \\ \frac{\partial u}{\partial \vec{n}} = 0 \quad \text{on} \quad \partial\Omega\times(0,T), \end{cases} \tag{3.3.3}$$

where this system is verified in the distributional sense. Moreover, the unique solution $u(x,t)$ is in $C^\infty((0,T)\times\overline{\Omega})$.

Proof. Firstly, we consider the proof of the existence of a solution, which is based on the Schauder fixed point argument [35, 53].

We introduce the solution space W of the problem (3.3.3) as follows:

$$W(0,T) = \left\{ w \in L^\infty(0,T;H^1(\Omega)), \quad \frac{dw}{dt} \in L^2(0,T;(H^1(\Omega))') \right\},$$

where $(H^1(\Omega))'$ is the dual space of $H^1(\Omega)$.

Let $w \in W$ such that

$$||w||_{L^\infty(0,T;L^2(\Omega))} \leq ||u_0||_{L^2(\Omega)}, \quad \left|\left|\frac{\partial w}{\partial t}\right|\right|_{L^2(0,T;(H^1(\Omega))')} \leq ||u_0||_{H^1(\Omega)}. \tag{3.3.4}$$

We consider the following linear problem $[E_w]$:

$$\left\langle \frac{\partial u(t)}{\partial t}, v \right\rangle_{(H^1(\Omega))' \times H^1(\Omega)} + \int_\Omega g_1(|\nabla G_\sigma * w(t)|)\nabla u(t)\nabla v(t)dx$$

$$+ \int_\Omega g_2(|\nabla G_\sigma * w(t)|)\nabla u(t)\nabla v(t)dx = 0$$

for all $v \in H^1(\Omega)$, a.e. $t \in [0,T]$. Since w and $\partial w/\partial t$ satisfy (3.3.4), then $|\nabla G_\sigma * w|$ and $|\nabla G_\sigma * (\frac{\partial w}{\partial t})|$ belong to $L^\infty((0,T);C^\infty(\Omega))$ and there exists a constant $M = M(G_\sigma, ||u_0||_{H^1(\Omega)})$ such that $|\nabla G_\sigma * w| \leq M$ and $|\nabla G_\sigma * (\frac{\partial w}{\partial t})| \leq M$ a.e. t, for all $x \in \Omega$. Since $g_1(s)$ and $g_2(s)$ are decreasing and positive, it follows that a.e. in $(0,T) \times \Omega$:

$$0 < \lambda_1 \leq g_1(|\nabla G_\sigma * w|) \quad \text{and} \quad 0 < \lambda_2 \leq g_2(|\nabla G_\sigma * w|).$$

By classical results on the parabolic equations [12, 28], the problem $[E_w]$ has a unique solution $U_w \in W$ [1, 12], satisfying the estimates

$$||U_w||_{L^\infty(0,T;H^1(\Omega))} \leq c_1, \tag{3.3.5}$$

$$||U_w||_{L^\infty(0,T;L^2(\Omega))} \leq ||u_0||_{L^2(\Omega)}, \tag{3.3.6}$$

$$\left|\left|\frac{\partial U_w}{\partial t}\right|\right|_{L^2(0,T;(H^1(\Omega))')} \leq ||u_0||_{H^1(\Omega)}, \tag{3.3.7}$$

where c_1 is a constant which only depends on G_σ, g_1, g_2 and $||u_0||_{H^1(\Omega)}$. Choosing $v = U_w$ in E_w, integrating over the interval $(0,t)$, we arrive to the inequality

$$\frac{1}{2}\int_\Omega U_w^2 dx + (\lambda_1 + \lambda_2)\int_0^t \int_\Omega |\nabla U_w|^2 dx\ ds \leq \frac{1}{2}\int_\Omega u_0^2\ dx, \tag{3.3.8}$$

which implies (3.3.6). Choosing $v = \partial U_w/\partial t$ in E_w, integrating by parts yields

$$\int_\Omega \left(\frac{\partial U_w}{\partial t}\right)^2 dx + \frac{1}{2}\int_\Omega g_1(|\nabla G_\sigma * w|)\frac{\partial|\nabla U_w|^2}{\partial t}\ dx + \frac{1}{2}\int_\Omega g_2(|\nabla G_\sigma * w|)\frac{\partial|\nabla U_w|^2}{\partial t}\ dx. \tag{3.3.9}$$

Integrating over the interval $(0,t)$ we arrive to that

$$\begin{aligned}\int_0^t\int_\Omega \left(\frac{\partial U_w}{\partial s}\right)^2 dx\ ds + \frac{1}{2}\int_\Omega g_1(|\nabla G_\sigma * w|)|\nabla U_w|^2\ dx + \frac{1}{2}\int_\Omega g_2(|\nabla G_\sigma * w|)|\nabla U_w|^2\ dx \\ = \frac{1}{2}\int_\Omega g_1(|\nabla G_\sigma * w|)|\nabla u_0|^2\ dx + \frac{1}{2}\int_0^t\int_\Omega g_1'(|\nabla G_\sigma * w|) \\ \times \nabla\left(G_\sigma * \frac{\partial w}{\partial s}\right)|\nabla U_w|^2\ dx\ ds + \frac{1}{2}\int_\Omega g_2(|\nabla G_\sigma * w|)|\nabla u_0|^2\ dx \\ + \frac{1}{2}\int_0^t\int_\Omega g_2'(|\nabla G_\sigma * w|) \times \nabla\left(G_\sigma * \frac{\partial w}{\partial s}\right)|\nabla U_w|^2\ dx\ ds.\end{aligned} \tag{3.3.10}$$

When $u \in W$ and (3.3.8), noticing that $|g_1(s)| \le k_1$ and $|g_2(s)| \le k_2$, we can deduce that

$$\begin{aligned}\int_0^t\int_\Omega \left(\frac{\partial U_w}{\partial s}\right)^2 dx\ ds + \frac{(\lambda_1+\lambda_2)}{2}\int_\Omega |\nabla u|^2\ dx \\ \le \int_\Omega |\nabla u_0|^2\ dx + \frac{M(k_1+k_2)}{4(\lambda_1+\lambda_2)}\int_\Omega u_0^2\ dx\ ds.\end{aligned} \tag{3.3.11}$$

Since $||u_0||_{H^1(\Omega)}$ is small, letting $(M(k_1+k_2)/4(\lambda_1+\lambda_2)) \le 1$ yields (3.3.5) and (3.3.7). From (3.3.5)-(3.3.7), we introduce the subspace W_0 of $W(0,T)$ defined by

$$\begin{aligned}W_0 = \{w \in W(0,T),\quad & w(0) = u_0, \\ & ||w||_{L^\infty(0,T;H^1(\Omega))} \le c_1,\quad ||w||_{L^\infty(0,T;L^2(\Omega))} \le ||u_0||_{L^2(\Omega)}, \\ & \left\|\frac{dw}{dt}\right\|_{L^2(0,T;(H^1(\Omega))')} \le ||u_0||_{H^1(\Omega)}\}.\end{aligned} \tag{3.3.12}$$

By construction, $w \to f(w) \equiv U_w$ is a mapping from W_0 into W_0. Moreover, W_0 is a not empty, convex and weakly compact in $W(0,T)$.

In order to apply the Schauder fixed point theorem, we need to prove that the mapping $f : w \to U_w$ is weakly continuous from W_0 into W_0. Let $\{w_j\}$ be a sequence in W_0 which converges weakly to some w in W_0 and let $u_j = U_{w_j}$. The sequence $\{w_j\}$

of W_0 contains a subsequence $\{w_j\}$ such that

$$\frac{du_j}{dt} \to \frac{du}{dt} \text{ weakly in } L^2(0,T;(H^1(\Omega))'),$$

$$u_j \to u \text{ in } L^\infty(0,T;L^2(\Omega)),$$

$$\frac{\partial u_j}{\partial x_i} \to \frac{\partial u}{\partial x_i} \text{ weakly in } L^\infty(0,T;L^2(\Omega)),$$

$$w_j \to w \text{ in } L^\infty(0,T;(L^2(\Omega)),$$

$$\frac{\partial G_\sigma}{\partial x_i} * w_j \to \frac{\partial G_\sigma}{\partial x_i} * w \text{ in } L^2(\Omega), \text{ a.e on } (0,T)\times\Omega,$$

$$g_1(\nabla G_\sigma * w_j) \to g_1(\nabla G_\sigma * w) \text{ in } L^2(0,T;L^2(\Omega)),$$

$$g_2(\nabla G_\sigma * w_j) \to g_2(\nabla G_\sigma * w) \text{ in } L^2(0,T;L^2(\Omega)),$$

$$u_j(0) \to u(0) \text{ in } L^2(\Omega).$$

Passing to the limit in the relation

$$\left\langle \frac{du_j(t)}{dt}, v \right\rangle + \int_\Omega g_1(|\nabla G_\sigma * w_j(t)|)\nabla u_j(t)\nabla v(t)dx + \int_\Omega g_2(|\nabla G_\sigma * w_j(t)|)\nabla u_j(t)\nabla v(t)dx = 0,$$

The above convergence allows us to pass to the limit in the problem (E_{w_j}) and obtain $u = U_w = f(w)$. Moreover, since the solution is unique, the whole sequence $u_j = f(w_j)$ converges weakly in W_0 to $u = f(w)$; that is, f is weakly continuous.

Second, the regularity of the solution using the general theory of parabolic equations and the bootstrap argument [13, 40], we can deduce that u is a strong solution of (3.3.3) and $u \in C^\infty((0,T)\times\Omega)$.

Finally, we shall proof that the uniqueness, following the idea in [28]. Let u_1 and u_2 be two solutions of (3.3.3). For almost every t in $[0,T]$, we have

$$\begin{cases} \frac{d}{dt}(u_1-u_2)(t) - \nabla\cdot(\beta_1(t)\nabla(u_1-u_2)(t)) - \nabla\cdot(\beta_2(t)\nabla(u_1-u_2)(t)) \\ = \nabla\cdot((\alpha_1-\beta_1)(t)\nabla u_1(t)) + \nabla\cdot((\alpha_2-\beta_2)(t)\nabla u_1(t)) \text{ in } \Omega, \\ \\ \frac{\partial(u_1-u_2)}{\partial\vec{n}} = 0 \text{ on } \partial\Omega\times(0,T), \end{cases} \tag{3.3.13}$$

in the distribution sense, where $\alpha_1(t)=g_1(|(\nabla G_\sigma * u_1)(t)|)$, $\alpha_2(t)=g_2(|(\nabla G_\sigma * u_1)(t)|)$,

$\beta_1(t)$=$g_1(|(\nabla G_\sigma * u_2)(t)|)$ and $\beta_2(t)$=$g_2(|(\nabla G_\sigma * u_2)(t)|)$.

Then multiplying the above equality by $v(t)$, integrating over Ω, and using the Neumann boundary conditions, we get a.e. $t \in [0,T]$,

$$\frac{1}{2}\frac{d}{dt}\int_\Omega (u_1-u_2)(t))v(t)+\int_\Omega \beta_1(t)\nabla v(t)\cdot(\nabla u_1(t)-\nabla u_2(t))+\int_\Omega \beta_2(t)\nabla v(t)\cdot(\nabla u_1(t)-\nabla u_2(t))$$

$$= -\int_\Omega (\alpha_1(t)-\beta_1(t))\nabla u_1(t)\cdot\nabla v(t) - \int_\Omega (\alpha_2(t)-\beta_2(t))\nabla u_1(t)\cdot\nabla v(t).$$

Taking $v(t) = u_1(t) - u_2(t)$ and using the bounds λ_1, λ_2 of g_1, g_2 respectively, we get

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}&||u_1(t)-u_2(t)||^2_{L^2(\Omega)} + \lambda_1||\nabla u_1(t)-\nabla u_2(t)||^2_{L^2(\Omega)} + \lambda_2||\nabla u_1(t)-\nabla u_2(t)||^2_{L^2(\Omega)}\\ &\le ||\alpha_1(t)-\beta_1(t)||_{L^\infty(\Omega)}||\nabla u_1(t)||_{L^2(\Omega)}||\nabla u_1(t)-\nabla u_2(t)||_{L^2(\Omega)}\\ &\quad + ||\alpha_2(t)-\beta_2(t)||_{L^\infty(\Omega)}||\nabla u_1(t)||_{L^2(\Omega)}||\nabla u_1(t)-\nabla u_2(t)||_{L^2(\Omega)}.\end{aligned} \tag{3.3.14}$$

Moreover, since g_1, g_2 and G_σ are smooth, we have

$$||\alpha_1(t)-\beta_1(t)||_{L^\infty(\Omega)} \le C_3||u_1(t)-u_2(t)||_{L^2(\Omega)},$$

$$||\alpha_2(t)-\beta_2(t)||_{L^\infty(\Omega)} \le C_4||u_1(t)-u_2(t)||_{L^2(\Omega)},$$

where C_3 and C_4 are constants which depends only on g_1, g_2, λ_1, λ_2 and G_σ. Combining these inequalities and using Schwarz inequality, we obtain

$$\frac{d}{dt}||u_1(t)-u_2(t)||^2_{L^2(\Omega)} \le C||\nabla u_1(t)||^2_{L^2(\Omega)}||u_1(t)-u_2(t)||^2_{L^2(\Omega)}, \tag{3.3.15}$$

for a.e., $0 \le t \le T$, where C is a constant which depends only on g_i, $i = 1, 2$, G_σ and u_0.

Since $u_1(0) - u_2(0) = u_0$, using Gronwall's inequality yields [28]

$$||u_1(t)-u_2(t)||^2_{L^2(\Omega)} \le 0;$$

that is, $u_1 = u_2$.

3.4 Some properties of weak solution

This section deal with two stages:

(a) Investigate the continuity with respect to initial data of the weak solution for (3.3.3).

(b) Investigate the stability of weak solution and the maximum principle. According to the uniqueness proof in Theorem 3.3.1.

Theorem 3.4.1. *Assume u is the weak solutions of problem (3.3.3) with the initial data u_0. Then*

$$\begin{aligned} &\int_\Omega (u-u_0)dx = 0, \\ &||u(.,t)-u_\Omega||_{L^2(\Omega)} \leq e^{\frac{-(\lambda_1+\lambda_2)t}{\mu}} ||u_0-u_\Omega||_{L^2(\Omega)}, \end{aligned} \tag{3.4.1}$$

a.e. $t \in [0,\infty)$, where $u_\Omega = \left(\frac{1}{|\Omega|}\right)\int_\Omega u_0\, dx$, and $|\Omega|$ is Lebesgue measure of Ω.

Proof. Let u be the solutions for problem (3.3.3) with the initial data u_0. For almost every t in [0,T], we have

$$\begin{cases} \frac{\partial u}{\partial t} = \nabla\cdot(g_1(|\nabla G_\sigma * u|)\nabla u) + \nabla\cdot(g_2(|\nabla G_\sigma * u|)\nabla u) = 0 \text{ on } \Omega\times(0,T), \\ u(x,0) = u_0(x) \ \text{ in } \ \Omega, \\ \frac{\partial u}{\partial \vec{n}} = 0 \text{ on } \ \partial\Omega\times(0,T), \end{cases} \tag{3.4.2}$$

in the distribution sense. Integrating over the interval $(0,T)$ and using the Neumann boundary conditions yield

$$\int_\Omega (u-u_0)dx = 0. \tag{3.4.3}$$

Then, multiplying the equation (3.4.2) by $(u-u_\Omega)$, and integrating by parts over Ω yields

$$\frac{1}{2}\frac{d}{dt}\int_\Omega (u-u_\Omega)^2 dx + \int_\Omega g_1(|\nabla G_\sigma * u|)|\nabla u|^2 dx + \int_\Omega g_2(|\nabla G_\sigma * u|)|\nabla u|^2 dx = 0. \tag{3.4.4}$$

Using the following Poincare-Wirtinger inequality [[4], page 148], we have

$$\left|\left|u - \frac{1}{|\Omega|}\int_\Omega u dx\right|\right|^2_{L^2(\Omega)} = ||u - u_\Omega||^2_{L^2(\Omega)} \leq \mu \int_\Omega |\nabla u|^2 dx, \tag{3.4.5}$$

with the constant $\mu \equiv \mu(\Omega)$. Substituting (3.4.5) to (3.4.4) yields

$$\frac{d}{dt}\int_\Omega (u - u_\Omega)^2 dx \leq -\frac{2(\lambda_1 + \lambda_2)}{\mu}\int_\Omega (u - u_\Omega)^2 dx. \tag{3.4.6}$$

Multiplying this inequality by $e^{\frac{2(\lambda_1+\lambda_2)t}{\mu}}$ and integrating over the interval (0,t) we arrive to the inequality

$$\int_\Omega (u - u_\Omega)^2 dx \leq e^{-\frac{2(\lambda_1+\lambda_2)t}{\mu}}\int_\Omega (u_0 - u_\Omega)^2 dx. \tag{3.4.7}$$

Hence, we obtain the assertion of the theorem.

Next, let us build upon the maximum principle as follows.

Theorem 3.4.2. *Let u be the weak solutions of problem (3.3.3) with the initial data $u_0 \in L^\infty(\Omega)$. Then*

$$\inf_{x\in\Omega} u_0 \leq u \leq \sup_{x\in\Omega} u_0. \tag{3.4.8}$$

Proof. Let $I := \sup\limits_{x\in\Omega} u_0$ and $j := \inf\limits_{x\in\Omega} u_0$. Multiply (3.3.3) by $(u - I)_+$, where

$$(u - I)_+ = \begin{cases} u - I, & \text{if } u - M > 0, \\ 0, & \text{otherwise,} \end{cases} \tag{3.4.9}$$

and integrate over Ω, we get

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}\int_\Omega (u(t) - I)^2_+ dx + \int_\Omega g_1(|\nabla G_\sigma * u_i|)|\nabla (u(t) - I)_+|^2 dx \\ &+ \int_\Omega g_2(|\nabla G_\sigma * u_i|)|\nabla (u(t) - I)_+|^2 dx = 0. \end{aligned} \tag{3.4.10}$$

Then

$$\frac{1}{2}\frac{d}{dt}\int_\Omega (u(t) - I)^2_+ dx \leq 0. \tag{3.4.11}$$

Therefore, $(1/2)(d/dt)\int_\Omega (u(t)-I)_+^2 dx$ is decreasing in t, and since

$$\int_\Omega (u(t)-I)_+^2 dx \geq 0, \qquad \int_\Omega (u(t)-I)_+^2 dx|_{t=0} = 0, \tag{3.4.12}$$

we have that

$$\int_\Omega (u(t)-I)_+^2 dx = 0, \quad \forall\ t \in [0,\infty), \tag{3.4.13}$$

and so

$$u(t) \leq \sup_{x\in\Omega} u_0 \ \text{ a. e. on } \ \Omega,\ \forall\ t>0. \tag{3.4.14}$$

Multiplying (3.3.3) by $(u-j)_-$, a similar argument yields that $u \geq j$ for all $t \in [0,\infty)$. Equation (3.4.8) is followed directly.

Now let us consider the following problem:

$$\begin{cases} \nabla\cdot(g_1(|\nabla G_\sigma * u|)\nabla u) + \nabla\cdot(g_2(|\nabla G_\sigma * u|)\nabla u) = 0 \ \text{ in } \ \Omega, \\ \frac{\partial u}{\partial \vec{n}} = 0, \quad \text{on } \ \partial\Omega, \\ \int_\Omega (u-u_0) = 0. \end{cases} \tag{3.4.15}$$

Theorem 3.4.3. *Assume $u_0 \in L^1(\Omega)$. Then the problem (3.4.15) admits a unique weak solution $u \in H^1(\Omega)$ such that*

$$\begin{cases} \int_\Omega g_1(|\nabla G_\sigma * u|)\nabla u\nabla v + \int_\Omega g_2(|\nabla G_\sigma * u|)\nabla u\nabla v = 0 \ \ \forall\ v \in C^\infty(\overline{\Omega}), \\ \int_\Omega (u-u_0) = 0. \end{cases} \tag{3.4.16}$$

Proof. It is clear that $u = u_\Omega$ is a unique solution for the problem (3.4.15).

Next we turn to the proof of the uniqueness of the solution for the problem (3.4.15). Let u_1 and u_2 be two weak solutions (3.4.15). Multiplying (3.4.15) by u, integrating over Ω, and using the Neumann boundary conditions, we get

$$\int_\Omega g_1(|\nabla G_\sigma * u|)|\nabla u|^2 + \int_\Omega g_2(|\nabla G_\sigma * u|)|\nabla u|^2 = 0. \tag{3.4.17}$$

Using the following Poincaré-Wirtinger inequality, we have

$$\left|\left|u - \frac{1}{|\Omega|}\int_\Omega u dx\right|\right|^2_{L^2(\Omega)} = ||u - u_\Omega||^2_{L^2(\Omega)} \leq \mu \int_\Omega |\nabla u|^2 dx, \tag{3.4.18}$$

with the constant $\mu \equiv \mu(\Omega)$. Substituting (3.4.15) and (3.4.18) to (3.4.17) yields

$$\int_\Omega (u - u_\Omega)^2 dx = 0. \tag{3.4.19}$$

Then

$$\int_\Omega (u_1 - u_2)^2 dx \leq \int_\Omega (u_1 - u_\Omega)^2 dx + \int_\Omega (u_2 - u_\Omega)^2 dx = 0. \tag{3.4.20}$$

That is, $u_1 = u_2$.

3.5 Convergent iterative scheme

Theorem 3.5.1. *Let $u_0 \in H^1(\Omega)$. The sequence $\{u^n\}$ defined by solving the iterative scheme*

$$\begin{cases} \frac{\partial u^{n+1}(t)}{\partial t} = \nabla\cdot(g_1(|\nabla G_\sigma * u^n(t)|)\nabla u^{n+1}(t)) \\ \qquad\qquad + \nabla\cdot(g_2(|\nabla G_\sigma * u^n(t)|)\nabla u^{n+1}(t)) \quad \text{in } (0,T)\times\Omega, \\ u^{n+1}(x,0) = u_0 \quad \text{in } \Omega, \\ \frac{\partial u^{n+1}(t)}{\partial \vec{n}} = 0 \quad \text{on } \partial(0,T)\times\Omega, \end{cases} \tag{3.5.1}$$

converges in $C([0,T];L^2(\Omega))$ to the strong solution of (3.3.3).

Proof. We denote by $\alpha^n = g_1(|\nabla G_\sigma * u^n|)$ and $\beta^n = g_2(|\nabla G_\sigma * u^n|)$. The problem (3.5.1) has a unique solution u^{n+1} by a classical theory on parabolic equations [12, 28]. It is clear that

$$\alpha^n \geq g_1(|\nabla G_\sigma * u^n|_{L^\infty(\Omega)}), \quad \beta^n \geq g_2(|\nabla G_\sigma * u^n|_{L^\infty(\Omega)}) \text{ a.e on } (0,T)\times\Omega.$$

Now we show that the sequence $\{u^n\}$ converges in $C([0,T];L^2(\Omega))$ to u, the strong

solution of (3.3.3). We observe, from the estimate (3.3.15), that

$$\frac{d}{dt}||u^{n+1}(t) - u(t)||^2_{L^2(\Omega)} \leq a(t)||u^n(t) - u(t)||^2_{L^2(\Omega)}, \tag{3.5.2}$$

where $a(t) = C(||\nabla u(t)||^2_{L^2(\Omega)})$.

Moreover, we have

$$||u^0 - u(t)||^2_{L^2(\Omega)} \leq C_0 \quad \forall \ \ t \in [0, T],$$

where C_0 is a constant which only depends on $||u_0||_{H^1(\Omega)}$. Then Gronwall's inequality yields, for any $t \in [0, T]$:

$$||u^1(t) - u(t)||^2_{L^2(\Omega)} \leq C_0 \left(\int_0^T a(s)ds \right),$$

and, by iteration,

$$||u^{n+1}(t) - u(t)||^2_{L^2(\Omega)} \leq C_0 \frac{1}{(n+1)!} \left(\int_0^T a(s)ds \right)^{n+1},$$

which implies that the sequence $\{u^n\}$ converges in $C([0, T]; L^2(\Omega))$ to the strong solution of (3.3.3).

3.6 Numerical experiments

In this Section, we perform two numerical experiments in 2D. The original images Lena and Boat shown in figure 2.5.1 have 256×256 pixels and each pixels has a value in [0, 255], and the Gaussian white noise is added by the normal imnoise function imnoise (I,'Gaussian', M, σ^2) (i.e., the mean M and variance σ^2) in Matlab. We first scale the intensities of the images into the range between zero and one before we begin our experiments. In our tests, we will use the PSNR as a criteria for the quality of restoration:

$$\text{PSNR} = 10\log_{10} \left(\frac{R^2}{\frac{1}{mn} \sum_{i,j}^{n} (u(i, j) - u_{new}(i, j))^2} \right), \tag{3.6.1}$$

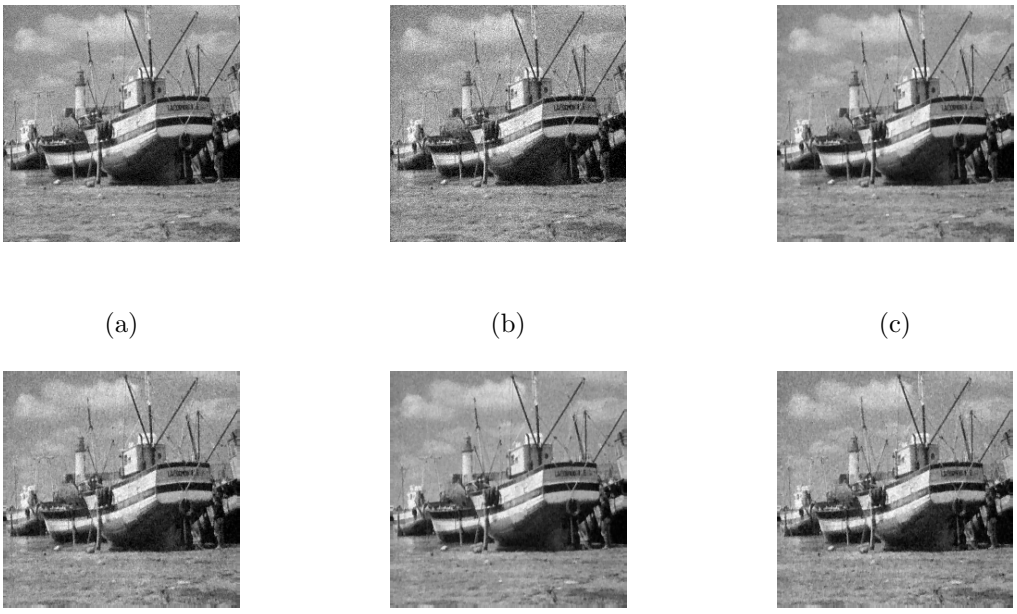
where $\{u(i, j) - u_{new}(i, j)\}$ are the differences of the pixel values between the restored and original images.

Now we demonstrate the numerical performance of our proposed second order model. The forward-backward difference scheme based on, see reference [82] with the Perona-Malik (PM) diffusivity $g_1(s) = \frac{1}{\sqrt{1+\frac{s^2}{\alpha^2}}}$ with $\alpha = 5$, see the references [18, 56, 82], and the Total variation (TV) diffusivity $g_2(s) = \frac{1}{s}$, in its regularized form $g_2(s) = \frac{1}{\sqrt{s^2+\epsilon^2}}$ with $\epsilon = 0.01$, is a popular choice, see the references [7, 16, 46]. It enforces piecewise constant results and therefore encourages sharp edges in the image. Throughout the experiments we have taken $\frac{\Delta t}{\Delta x^2} = 0.1$ [46] and Lagrange multiplier $\lambda = 0.85$ as in [15] and [17].

(a) (b) (c)

(d) (e) (f)

Figure 3.6.1: Noisy Lena images with different levels of Gaussian noise (a)- (b), σ^2= 0.002, 0.004, respectively; (c)-(d) corresponding denoised images by Catt*é* model with Perona-Malik (PM) diffusivity; (e)-(f) by our model (3.2.1).



(a) (b) (c)

(d) (e) (f)

Figure 3.6.2: Noisy Boat images with different levels of Gaussian noise (a)- (b), σ^2= 0.002, 0.004, respectively; (c)-(d) corresponding denoised images by Catté model with Perona-Malik (PM) diffusivity; (e)-(f) by our model (3.2.1).

Table 3.6.1: Results obtained by using models (3.2.1) and Catté et al. [13] applied to the images in Figure 3.6.1 with two different levels of Gaussian noise (σ^2= 0.002 and 0.004).

Images	PSNR	Images	PSNR (Catté)	Images	PSNR (3.2.1)
Fig. 3.6.1(a)	27.06	Fig. 3.6.1(c)	30.31	Fig. 3.6.1(e)	30.78
Fig. 3.6.1(b)	24.10	Fig. 3.6.1(d)	27.44	Fig. 3.6.1(f)	28.08
-	-	No. of iterations	1000	No. of iterations	1000

Table 3.6.2: Results obtained by using models (3.2.1) and Catté et al. [13] applied to the images in Figure 3.6.2 with two different levels of Gaussian noise (σ^2= 0.002 and 0.004).

Images	PSNR	Images	PSNR (Catté)	Images	PSNR (3.2.1)
Fig. 3.6.2(a)	27.03	Fig. 3.6.2(c)	29.74	Fig. 3.6.2(e)	30.13
Fig. 3.6.2(b)	24.07	Fig. 3.6.2(d)	27.12	Fig. 3.6.2(f)	27.64
-	-	No. of iterations	1000	No. of iterations	1000

3.7 Conclusion

We have presented a new second order partial differential equation based nonlinear diffusion model for image denoising. The main idea is to apply a priori smoothness on the solution image. The forward-backward difference schemes are used to discretize model (3.2.1) and Catté model. The model (3.2.1) gives larger PSNR values than that of Catté model, at the same iteration numbers.

Chapter 4

Efficient PDE-based nonlinear diffusion and Time dependent models for image denoising

4.1 Introduction

In this Chapter, we present a new nonlinear anisotropic diffusion model which incorporates adaptive information computed from the image at scale t. Following [3, 9], well-posedness of the proposed scheme is proved using the theory of viscosity solutions. We derive theoretical considerations for anisotropic diffusion model. We present proof of the viscosity solution of model (4.2.6). Besides, we propose a time dependent model for solving total variation (TV) minimization problem in image denoising. This is a constrained optimization type of numerical algorithm for removing noise from images. The constraints are imposed using Lagrange's multipliers and the solution is obtained using the gradient projection method. 1D and 2D numerical experimental results by explicit numerical schemes are discussed.

The nonlinear diffusion method for image denoising and edge detection was first introduced by Perona and Malik [56]. This method is based on a diffusion process governed by a partial differential equation (PDE), where diffusion amount depends on the gradient of images.

Mathematically, $u_0 : \Omega \to \mathbb{R}$ represents a noisy version of a true image, and it is obtained by the following imaging process

$$u_0(x, y) = u(x, y) + n(x, y), \tag{4.1.1}$$

where $u(x,y)$ denotes the desired clean image, $u_0(x,y)$ denotes the pixel values of a noisy image for $x,y \in \Omega$, $\Omega \subset \mathbb{R}^2$ is a bounded domain, usually a rectangle and $n(x,y)$ is additive white noise assumed to be close to Gaussian. The values $n(i,j)$ of n at the pixels (i,j) are independent random variables, each with a Gaussian distribution of zero mean and variance σ^2.

In our tests, we will use the peak signal to noise ratio (PSNR) as a criteria for the quality of restoration:

$$\text{PSNR} = 10\log_{10}\left(\frac{R^2}{\frac{1}{mn}\sum_{i,j}^{n}(u(i,j)-u_{new}(i,j))^2}\right), \tag{4.1.2}$$

where $\{u(i,j)-u_{new}(i,j)\}$ is the difference of the pixel values between the restored and original images.

The choice of the diffusivity c is very important in controlling the smoothing and even enhancement of edges. The Charbonnier diffusivity $c(s) = \frac{1}{\sqrt{1+(|s|^2/K^2)}}$, that is related to the convex regularizer $\psi(s^2) = \sqrt{K^4 + K^2 s^2} - K^2$, see references [18, 79], is used in our experiments.

4.2 Physical background for anisotropic diffusion model for image denoising

In general, variational deblurring and denoising of an image can be achieved by minimizing the energy functional presented in [82],

$$E(u) = \int_{\Omega} \psi(|\nabla u|^2)\, dx + \frac{\lambda}{2}\int_{\Omega}(u-u_0)^2\, dx. \tag{4.2.1}$$

The Euler-Lagrange equation associated with (4.2.1) with homogeneous Neumann boundary conditions is given by

$$\begin{cases} 0 = -\ \text{div}(\psi'(|\nabla u|^2)\nabla u) + \lambda\ (u-u_0), & x \in \Omega, \\ \\ \frac{\partial u}{\partial \vec{n}} = 0, & x \in \partial\Omega, \end{cases} \tag{4.2.2}$$

where $\partial\Omega$ is the boundary of Ω and $\vec{n}$ is the outward normal to $\partial\Omega$.

The resulting gradient descent equation is

$$u_t = \text{div}(c(|\nabla u|)\nabla u) - \lambda\ (u - u_0), \tag{4.2.3}$$

with $u(x, 0)$ given as initial data (the original noisy image $u_0(x)$ used as initial guess), homogeneous Neumann boundary conditions, i.e., $\frac{\partial u}{\partial \vec{n}} = 0$ on the boundary of the domain. It is also known as diffusion-reaction equation where the diffusion term with diffusivity $c(s) = \psi'(s^2)$ is related to the regulariser in the energy functional.

Applying a priori smoothness on the solution image, our nonlinear anisotropic diffusion model becomes,

$$u_t = \text{div}(c(|\nabla G_\sigma * u|)\nabla G_\sigma * u) - \lambda\ (G_\sigma * u - u_0). \tag{4.2.4}$$

Witkin [83] noticed that the convolution of the signal with Gaussians at each scale was equivalent to solving the heat equation with the signal as initial datum. The term $(G_\sigma * \nabla u)(x, t) = (\nabla G_\sigma * u)(x, t)$, which appears inside the divergence term of (4.2.4), is simply the gradient of the solution at time σ of the heat equation with $u(x, 0)$ as initial datum.

In order to preserve the notion of scale in the gradient estimate, it is convenient that this kernel G_σ depends on a scale parameter [45]. In fact, the function G_σ can be considered as "low-pass filter" or any smoothing kernel, i.e., a denoising technique is used before solving the nonlinear diffusion problem [3, 13].

We use the following class of functions for the diffusion equation, given in [8, 71],

$$c(x, |\nabla u|) = \alpha(x)c_g(|\nabla u|). \tag{4.2.5}$$

Here α is the adaptive parameter estimated at each pixel $x \in \Omega$. The function c_g depends on the gradient image $|\nabla u|$ and can be chosen similar to $c(s)$. If we choose $\alpha(x) = 1$, $c_g = c(s)$ and $G_\sigma * u$ as u then the model (4.2.4) can be written as:

$$\frac{\partial u}{\partial t} = \text{div}(c(x, |\nabla u|)\nabla u) - \lambda(u - u_0). \tag{4.2.6}$$

4.3 Theoretical considerations

In this Section, motivated by Alvarez et al. [3], we want to present the viscosity solution for model (4.2.6).

$$\begin{aligned} &\frac{\partial u}{\partial t} = \operatorname{div}(c(x,|\nabla u|)\nabla u) - \lambda(u-u_0), \quad x \in \mathbb{R}^n,\ t \in \mathbb{R}_+, \\ &u(x,0) = u_0(x), \quad x \in \mathbb{R}^n. \end{aligned} \tag{4.3.1}$$

Let us first introduce two auxiliary functions depending on x and p from $\mathbb{R}^n$, a symmetric matrix-valued one a and a vector one χ. We denote

$$a_{ij}(x,p) = c(x,|p|)\delta_{ij} + c_y(x,|p|)\frac{p_i p_j}{|p|}, \tag{4.3.2}$$

$$\chi_i(x,p) = \frac{\partial c(x,|p|)}{\partial x_i}. \tag{4.3.3}$$

Here δ_{ij} is Kronecker's delta, and c_y is the partial derivative of $c(x,y)$ with respect to the second variable.

Motivated by Alvarez et al. [2], we consider the case of spatially periodic boundary conditions. We will assume that there is an orthogonal basis b_i in $\mathbb{R}^n$ so that

$$u(.,x+b_i) = u(.,x), \quad x \in \mathbb{R}^n, \quad i = 1,2,....n. \tag{4.3.4}$$

Let u_0 is Lipschitz and satisfies (4.3.4). Of course, c (and thus a and χ) should also satisfy the same spatial periodicity restriction (with respect to x but not to y or p). Functions a and χ are continuous, bounded, periodic and continuously differentiable in x and their x-derivatives are uniformly (w.r.t. p) bounded,

$$a_{ij}(x,p)\xi_i\xi_j \geq C\left[\operatorname{mod}\left(\frac{\partial a(x,p)}{\partial x_k}\right)\right]_{ij}\xi_i\xi_j, \quad k = 1,.....n,\ \xi,x,p \in \mathbb{R}^n. \tag{4.3.5}$$

Here $\lambda \geq 0$ and below C stands for a generic positive constant, which can take different values in different lines.

We first recall the definition of viscosity sub-/supersolution of (4.3.1), if for any $\phi \in C^2([0,T]\times\mathbb{R}^n)$ and any point $(x_0,t_0) \in (0,T]\times\mathbb{R}^n$, at which $u-\phi$ attains local

maximum/minimum [29].

$$\frac{\partial\phi(x_0,t_0)}{\partial t} - \mathrm{div}(c(x_0,|\nabla\phi(x_0,t_0)|)\nabla\phi(x_0,t_0)) + \lambda(u(x_0,t_0) - u_0(x_0)) \leq 0/ \geq 0. \quad (4.3.6)$$

A viscosity solution is a function which is both a subsolution and a supersolution.

Lemma 4.3.1. *Let A and B be quadratic matrices of order n. Assume that B is symmetric and there is a constant $M \geq 0$ such that*

$$MA_{ij}\xi_i\xi_j \geq mod\ (B)_{ij}\xi_i\xi_j, \quad \forall\ \xi \in \mathbb{R}^n. \quad (4.3.7)$$

Then for any matrix U (of the same order but not necessarily symmetric) one has

$$Tr^2(BU^\top) \leq M||B||Tr(UAU^\top), \quad (4.3.8)$$

where $||.||$ denotes the operator norm of a matrix and mod (B) be the matrix whose entries are the absolute values of the entries of B.

Proof. Formula (4.3.7) and (4.3.8) are invariant with respect to orthogonal changes of bases. Thus, without loss of generality, we may assume that B is already diagonalized by an orthogonal transform. Then

$$Tr^2(BU^\top) = (B_{ii}U_{ii})^2 \leq ||B|||B_{ii}U_{ii}^2$$

$$= ||B||(\mathrm{mod}(B)_{ii}U_{ii}^2 \leq ||B||(\mathrm{mod}(B)_{ii}U_{ki}U_{kj}$$

$$= ||B||(\mathrm{mod}(B)_{ij}U_{ki}U_{kj} \leq M||B||A_{ij}U_{ki}U_{kj} = M||B||Tr(UAU^\top).$$

Theorem 4.3.1. *The problem (4.3.1) has a unique viscosity solution u in $C([0,T]\times\mathbb{R}^n)\cap L^\infty(0,T,W^{1,\infty}(\mathbb{R}^n))$ for any $T\in[0,\infty)$, provided that u_0 is Lipschitz continuous on $\mathbb{R}^n$, and if $v \in C(\mathbb{R}^n\times[0,T))$ is a viscosity solution of (4.3.1) with u_0 replaced by a Lipschitz continuous function v_0, then for all $T\in[0,\infty)$, there exists a constant $C>0$, depending only on u_0, v_0 and T, such that*

$$\sup_{0\leq t\leq T}||u(x,t)-v(x,t)||_{L^\infty(\mathbb{R}^n)} \leq C||u_0-v_0||_{L^\infty(\mathbb{R}^n)}. \quad (4.3.9)$$

Moreover, $\inf\limits_{\mathbb{R}^n} u_0 \leq u(x,t) \leq \sup\limits_{\mathbb{R}^n} u_0$.

Proof. If u is a viscosity solution of equation (4.3.1) on $\mathbb{R}^n \times \mathbb{R}_+$, then

$$\inf_{\mathbb{R}^n} u_0 \le u(x,t) \le \sup_{\mathbb{R}^n} u_0, \quad \text{on} \quad \mathbb{R}^n \times \mathbb{R}_+. \tag{4.3.10}$$

Let $\phi(x,t) = \delta t$, then, at the point (x_0, t_0), $t_0 > 0$, of the global maximum of $u(x,t) - \delta t$, (4.3.6) gives $\delta + \lambda(u(t_0, x_0) - u_0(x_0)) \le 0$, when $u(x_0, t_0) < u_0(x_0)$, so we get a contradiction since $u(x_0, t_0) - \delta t_0 \ge u_0(x_0)$ due to the fact that (x_0, t_0) is the global maximum point of $u(x,t) - \delta t$; thus the function $u(x,t) - \delta t$ attains its global maximum at $t = 0$, and it remains to let $\delta \to 0^+$, we get (4.3.10).

Now, we establish a formal a priori estimate for $\sup_{\mathbb{R}^n} |\nabla u|$. Observe that (4.3.1) is equivalent to

$$\frac{\partial u}{\partial t} = [a_{ij}(x, \nabla u) u_{x_i x_j} + \chi_i(x, \nabla u) u_{x_i}] - \lambda(u - u_0). \tag{4.3.11}$$

Differentiating (4.3.11) with respect to each x_k, $k = 1, ..., n$, multiplying by $2u_{x_k}$ and taking a summation w.r.t. k, we get

$$\begin{aligned} \gamma(|\nabla u|^2) &:= \frac{\partial |\nabla u|^2}{\partial t} - a_{ij}(x, \nabla u) \frac{\partial^2}{\partial x_i \partial x_j} |\nabla u|^2 - \frac{\partial a_{ij}(x, \nabla u)}{\partial p_l} u_{x_i x_j} \frac{\partial}{\partial x_l} |\nabla u|^2 \\ &- \chi_i(x, \nabla u) \frac{\partial}{\partial_i} |\nabla u|^2 - \frac{\partial \chi_i(x, \nabla u)}{\partial p_l} u_{x_i} \frac{\partial}{\partial x_l} |\nabla u|^2 + 2\lambda(u_{x_k} - (u_0)_{x_k}) u_{x_k} \\ &= -2a_{ij}(x, \nabla u) u_{x_k x_i} u_{x_k x_j} + 2 \frac{\partial a_{ij}(x, \nabla u)}{\partial x_k} u_{x_i x_j} u_{x_k} + 2 \frac{\partial \chi_{ij}(x, \nabla u)}{\partial x_k} u_{x_i} u_{x_k}. \end{aligned} \tag{4.3.12}$$

The Lemma 4.3.1 gives opportunity to discharge the undesired influence of the second term in the right-hand side of (4.3.12). For the second term, due to the Lemma 4.3.1 and Cauchy's inequality, we have

$$\begin{aligned} \left| 2 \frac{\partial a_{ij}(x, \nabla u)}{\partial x_k} u_{x_i x_j} u_{x_k} \right| &\le C |u_{x_k}| \sqrt{a_{ij}(x, \nabla u) u_{x_k x_i} u_{x_k x_j}} \\ &\le a_{ij}(x, \nabla) u_{x_k x_i} u_{x_k x_j} + C |\nabla u|^2. \end{aligned} \tag{4.3.13}$$

The sum of the absolute values of the subsequent terms of the right-hand side of (4.3.12) does not exceed $C(1 + |\nabla u|^2)$. Thus,

$$\gamma(|\nabla u|^2) \le C(1 + |\nabla u|^2), \tag{4.3.14}$$

so

$$\gamma(e^{-Ct}(1+|\nabla u|^2)) \leq 0. \tag{4.3.15}$$

From the weak maximum principle for the weakly parabolic operator γ one easily concludes that

$$|\nabla u|^2 \leq C. \tag{4.3.16}$$

Using (4.3.10) and (4.3.16), by means of the approach from [2] we can get the uniform Holder estimate

$$|u(x,t) - y(x,s)|^2 \leq C|t-s|. \tag{4.3.17}$$

From (4.3.10), (4.3.16) and (4.3.17), the solutions of these problems are uniformly bounded and equicontinuous on $\mathbb{R}^n \times [0,T]$. Then we can select a uniformly converging sequence of approximate solutions, and pass to the limit in the viscosity sense using the general consistency/stability results from [24]. The uniqueness of solutions follows from the stability estimate (4.3.9). This bound may be shown by revisiting the proof of a similar bound in [66]. We only point out that the matrix τ [66] is replaced by

$$\tau = \begin{pmatrix} D_1 & \sqrt{D_1}\sqrt{D_2} \\ \sqrt{D_1}\sqrt{D_2} & D_2 \end{pmatrix}, \tag{4.3.18}$$

where

$$D_1 = a\left(x_0, \frac{|x_0 - y_0|^2(x_0 - y_0)}{\delta}\right), \quad D_2 = a\left(y_0, \frac{|x_0 - y_0|^2(x_0 - y_0)}{\delta}\right),$$

and the notation within is taken from [66]. note that the $2n \times 2n$ matrix τ is symmetric and positive-semidefinite.

4.4 Numerical experiments for anisotropic diffusion model

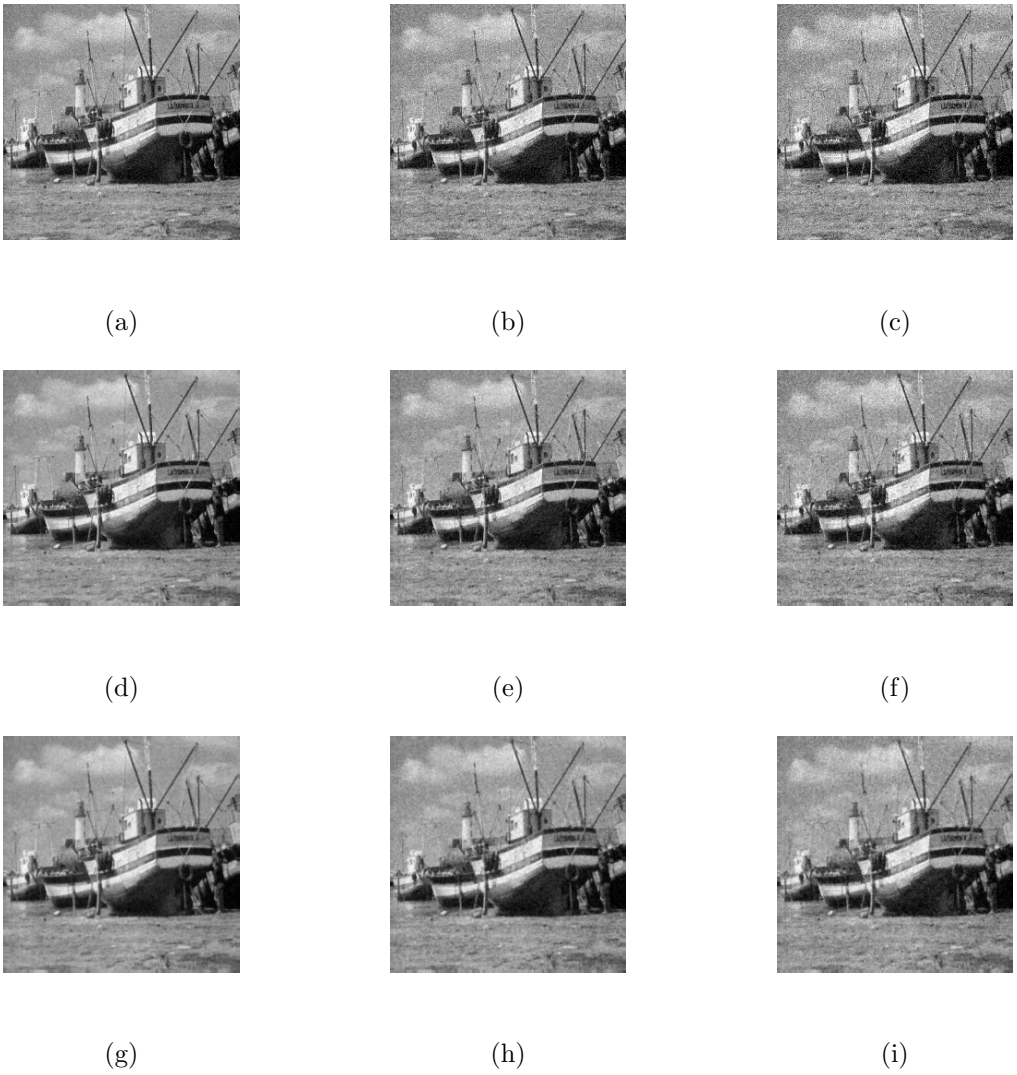
We have used two gray scale images as shown in Figure 2.5.1. The pixel values of all images lie in interval [0, 255]. The Gaussian white noise is added by the normal imnoise function imnoise (I,'Gaussian', M, σ^2), i.e., the mean M and variance σ^2 in Matlab. We first scale the intensities of the images into the range between zero and one before we begin our experiments. We have taken $\Delta t/\Delta x^2 = 0.4$, see [41], Charbonnier diffusivity $K = 5$ and Lagrange multiplier $= 0.85$ as in [15] and [17] in our all experiments.

(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 4.4.1: (top row) Noisy Lena images with different levels of Gaussian noise (a)-(c), $\sigma^2 = 0.002,\ 0.004,\ 0.006$, respectively; (second row) (d)-(f) corresponding denoised images by model (4.2.3); (third row) (g)-(i) by model (4.2.4).



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 4.4.2: (top row) Noisy Boat images with different levels of Gaussian noise (a)-(c), $\sigma^2 = 0.002,\ 0.004,\ 0.006$, respectively; (second row) (d)-(f) corresponding denoised images by model (4.2.3); (third row) (g)-(i) by model (4.2.4).

Table 4.4.1: Results obtained by using models (4.2.3) and (4.2.4) applied to the images in Figures 4.4.1(a) and 4.4.2(a) with Gaussian white noise ($\sigma^2 = 0.002$).

Images	PSNR Noisy Image	Images	PSNR (4.2.3)	Images	PSNR (4.2.4)
Fig. 4.4.1(a)	27.02	Fig. 4.4.1(d)	30.33	Fig. 4.4.1(g)	30.48
Fig. 4.4.2(a)	27.05	Fig. 4.4.2(d)	29.75	Fig. 4.4.2(g)	30.01
-	-	No. of iterations	400	No. of iterations	200

Table 4.4.2: Results obtained by using models (4.2.3) and (4.2.4) applied to the images in Figures 4.4.1(b) and 4.4.2(b) with Gaussian white noise (σ^2 =0.004).

Images	PSNR Noisy Image	Images	PSNR (4.2.3)	Images	PSNR (4.2.4)
Fig. 4.4.1(b)	24.08	Fig. 4.4.1(e)	27.14	Fig. 4.4.1(h)	28.78
Fig. 4.4.2(b)	24.09	Fig. 4.4.2(e)	26.80	Fig. 4.4.2(h)	28.45
-	-	No. of iterations	400	No. of iterations	200

Table 4.4.3: Results obtained by using models (4.2.3) and (4.2.4) applied to the images in Figures 4.4.1(c) and 4.4.2(c) with Gaussian white noise (σ^2= 0.006).

Images	PSNR Noisy Image	Images	PSNR (4.2.3)	Images	PSNR (4.2.4)
Fig. 4.4.1(c)	22.36	Fig. 4.4.1(f)	25.17	Fig. 4.4.1(i)	27.46
Fig. 4.4.2(c)	22.33	Fig. 4.4.2(f)	24.96	Fig. 4.4.2(i)	27.28
-	-	No. of iterations	400	No. of iterations	200

4.5 Time dependent model for 2D

Most conventional variational methods involve a least squares L^2 fit because this leads to linear equations. The first attempt along these lines was made by Phillips [57] and later refined by Twomey et al. [73, 74] in one-dimensional case.

The total variation based image denoising model, which is based on the constrained minimization problem appeared in [61], is as follows:

$$minimize \int_{\Omega} |\nabla u| \; dx \; dy = \int_{\Omega} \sqrt{u_x^2 + u_y^2} \; dx \; dy, \tag{4.5.1}$$

subject to constraints

$$\int_{\Omega} u \; dx \; dy = \int_{\Omega} u_0 \; dx \; dy, \tag{4.5.2}$$

and

$$\int_{\Omega} \frac{1}{2}(u - u_0)^2 \; dx \; dy = \sigma^2. \tag{4.5.3}$$

The first constraint corresponds to the assumption that the noise has zero mean, and the second constraint uses a priori information that the standard deviation of the noise $n(x, y)$ is σ.

The Euler-Lagrange equation is given by,

$$0 = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) - \lambda_1 - \lambda_2(u - u_0), \tag{4.5.4}$$

in Ω, with $\frac{\partial u}{\partial n} = 0$ on the boundary of the domain.

Since (4.5.4) is not well defined at points where $\nabla u = 0$, due to the presence of the term $\frac{1}{|\nabla u|}$, it is common to slightly perturb the TV algorithm to become

$$\int_{\Omega} |\nabla u|_{\beta} \; dx \; dy = \int_{\Omega} \sqrt{u_x^2 + u_y^2 + \beta} \; dx \; dy, \tag{4.5.5}$$

where β is a small positive parameter [17].

The solution procedure uses a parabolic equation with time as an evolution parameter, or equivalently, the gradient descent method. This means that we solve

$$u_t = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) - \lambda(u - u_0), \tag{4.5.6}$$

for $t > 0,\ x, y \in \Omega$ with $u(x, y, 0)$ given as initial data and $\frac{\partial u}{\partial n} = 0$ on the boundary of the domain.

Applying a priori smoothness on the solution image, our new time dependent model becomes,

$$u_t = \nabla \cdot \left(\frac{\nabla G_\sigma * u}{|\nabla G_\sigma * u|} \right) - \lambda(G_\sigma * u - u_0), \tag{4.5.7}$$

for $t > 0,\ x, y \in \Omega$ with $u(x, y, 0)$ given as initial data and $\frac{\partial u}{\partial n} = 0$ on the boundary of the domain. It should be noticed that (4.5.7) only replaces u in (4.5.6) by its estimate $G_\sigma * u$.

Witkin [83] noticed that the convolution of the signal with Gaussians at each scale was equivalent to solving the heat equation with the signal as initial datum. The term $(G_\sigma * \nabla u)(x, y, t) = (\nabla G_\sigma * u)(x, y, t)$, which appears inside the divergence term of (4.5.7), is simply the gradient of the solution at time σ of the heat equation with $u(x, y, 0)$ as initial datum. In order to preserve the notion of scale in the gradient estimate, it is convenient that this kernel G_σ depends on a scale parameter [45]. In fact, the function G_σ can be considered as "low-pass filter" or any smoothing kernel, i.e., a denoising technique is used before solving the nonlinear diffusion problem [3, 13].

The first constraint (4.5.4) is dropped because it is automatically enforced by the evolution procedure, i.e., the mean of $u(x, y, 0)$ is the same as that of $u_0(x, y)$. As t increases, a denoised version of image is realised.

To compute $\lambda(t)$, we multiply (4.5.6) by $(u - u_0)$ and integrate by parts over Ω. If steady state has been reached, the left side of (4.5.6) vanishes. We then have,

$$\lambda = -\frac{1}{2\sigma^2} \int_\Omega \left[|\nabla u| - \left(\frac{(u_0)_x u_x}{|\nabla u|} + \frac{(u_0)_y u_y}{|\nabla u|} \right) \right] \, dx \, dy. \tag{4.5.8}$$

This gives us a dynamic value $\lambda(t)$, which appears to converge as $t \to \infty$. The theoretical justification for this approach comes from the fact that it is merely the gradient projection method of Rosen [59].

We still write $G_\sigma * u$ as u. Let u_{ij}^n be the approximation to the value $u(x_i, y_j, t_n)$, where

$$x_i = i\Delta x,\ y_j = j\Delta x,\ i, j = 1, 2, ..., N,$$

$$N\Delta x = 1, \quad t_n = n\Delta t, \quad n = 0, 1, ...,$$

$$u_{ij}^n = u(x_i, y_j, t_n),$$

$$u_{ij}^0 = u_0(i\Delta x, j\Delta x) + \sigma\phi(i\Delta x, j\Delta x). \tag{4.5.9}$$

The modified initial data are chosen so that the constraints are satisfied initially, i.e, ϕ has mean zero and L^2 norm one.

The explicit partial derivatives of model (4.5.6) and model (4.5.7) can be expressed as:

$$u_t = \frac{u_{xx}(u_y^2 + \beta) - 2u_{xy}u_x u_y + u_{yy}(u_x^2 + \beta)}{(u_x^2 + u_y^2 + \beta)^{\frac{3}{2}}} - \lambda(u - u_0). \tag{4.5.10}$$

We define the derivative terms as,

$$u_{ij}^x = \frac{u_{i+1,j}^n - u_{i-1,j}^n}{2\Delta x}; \qquad u_{ij}^y = \frac{u_{i,j+1}^n - u_{i,j-1}^n}{2\Delta x};$$

$$u_{ij}^{xx} = \frac{u_{i+1,j}^n - 2u_{i,j}^n + u_{i-1,j}^n}{\Delta x^2}; \qquad u_{ij}^{yy} = \frac{u_{i,j+1}^n - 2u_{i,j}^n + u_{i,j-1}^n}{\Delta x^2};$$

$$u_{ij}^{xy} = \frac{u_{i+1,j+1}^n - u_{i-1,j+1}^n - u_{i+1,j-1}^n + u_{i-1,j-1}^n}{4\Delta x \Delta x}; \qquad u_{ij}^t = \frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t}.$$

We let,

$$r_{ij}^n = u_{ij}^{xx}((u_{ij}^y)^2 + \beta) - 2u_{ij}^{xy}u_{ij}^x u_{ij}^y + u_{ij}^{yy}((u_{ij}^x)^2 + \beta), \tag{4.5.11}$$

and

$$p_{ij}^n = ((u_{ij}^x)^2 + (u_{ij}^y)^2 + \beta)^{\frac{3}{2}}. \tag{4.5.12}$$

Then (4.5.10) reads as follows:

$$u_{ij}^t = \frac{r_{ij}^n}{p_{ij}^n} - \lambda(u_{ij}^n - u_0(i\Delta x, j\Delta x), \tag{4.5.13}$$

with boundary conditions

$$u_{i,N}^n = u_{i,N-1}^n, \; u_{N,j}^n = u_{N-1,j}^n, \; u_{1,j}^n = u_{2,j}^n, \; u_{i,1}^n = u_{i,2}^n. \tag{4.5.14}$$

The explicit method is stable and convergent for $\frac{\Delta t}{\Delta x^2} \leq 0.5$, see [41].

4.6 Time dependent model for 1D

The 2D model described before is more regular than the corresponding 1D model because the 1D original optimization problem is barely convex. For the sake of understanding the numerical behavior of our schemes, we also discuss the 1D model. The Euler-Lagrange equation in the 1D case reads as follows:

$$0 = \left(\frac{u_x}{|u_x|}\right)_x - \lambda_1 - \lambda_2(u - u_0). \tag{4.6.1}$$

This equation can be written either as

$$0 = \left(\frac{u_x}{|u_x|_\beta}\right)_x - \lambda_1 - \lambda_2(u - u_0), \tag{4.6.2}$$

using the small regularizing parameter $\beta > 0$ introduced in [46], or

$$0 = \delta(u_x)u_{xx} - \lambda_1 - \lambda_2(u - u_0), \tag{4.6.3}$$

using the δ-function.

Our model in 1D will be

$$u_t = \frac{\beta}{(\beta + u_x^2)^{\frac{3}{2}}}u_{xx} - \lambda(u - u_0), \tag{4.6.4}$$

where $\beta > 0$ is small regularizing parameter. The parameter $\beta > 0$ in this model is estimated from the local amount of noise. We have found for our model, through our numerical experiments in 1D, that β can be estimated as the standard deviation of the noise.

We can also state our model in terms of the δ function as

$$u_t = \delta(u_x)u_{xx} - \lambda(u - u_0). \tag{4.6.5}$$

In this chapter, we approximate δ, see the reference [46], by

$$\delta(k) \approx \beta.(k^2 + \beta)^{-\frac{3}{2}}. \tag{4.6.6}$$

These evolution models are initialized with the noisy signal u_0, homogeneous Neu-

mann boundary conditions, and with a prescribed Lagrange multiplier for slightly noisy signals.

We have estimated $\lambda > 0$ near the maximum value such that the explicit scheme is stable under appropriate CFL ($\frac{\Delta t}{\Delta x^2} < 0.25$) restrictions [46], provided β is chosen to be the standard deviation of the noise.

The following is the explicit numerical scheme of model (4.6.4). Let u_i^n be the approximation to the value $u(x_i, t_n)$, where $x_i = i\Delta x$ and $t_n = n\Delta t,\ n \geq 1$. We define the derivative terms as,

$$u_x = \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x}; \quad u_{xx} = \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{\Delta x^2};$$

$$u_t = \frac{u_i^{n+1} - u_i^n}{\Delta t}.$$

We let,

$$b_i = \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x}. \tag{4.6.7}$$

Then (4.6.4) reads as follows:

$$u_i^{n+1} = u_i^n + \Delta t \left[\frac{\beta}{(\beta + b_i^2)^{\frac{3}{2}}} \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{\Delta x^2} \right] - \Delta t \lambda (u_i^n - u_0(x_i)). \tag{4.6.8}$$

4.7 Numerical experiments for time dependent model for 1D

We, as an example, have taken 1D signals $u(x) = \exp(0.1x).|sin(x)|,\ x \in [0, 25]$ and $u(x) = |sin(x) + cos(x)|,\ x \in [0, 25]$ given in figures 4.7.1(a) and 4.7.1(b) respectively. When Gaussian white noise is added to them, we get noisy signals.

In our test, we will use the signal to noise ratio (SNR) of the signal u to measure the level of noise, defined as

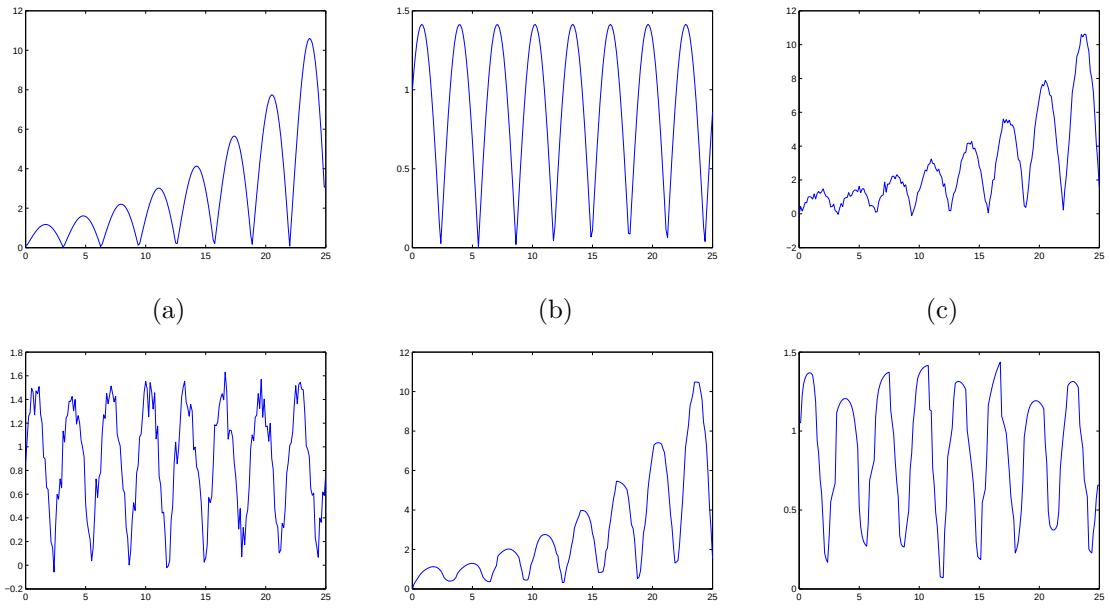
$$\text{SNR} = \frac{||u - \overline{u}||_{L^2}}{\sigma}, \tag{4.7.1}$$

where $\overline{u}$ is the mean of the signal u, i.e., the ratio of the standard deviation of the signal over the standard deviation of the noise.

The standard deviation of noisy signals (given in Figures 4.7.1(c) and 4.7.1(d)) are approximately $\sigma \approx 0.5$ and $\sigma \approx 0.04$ respectively whereas their SNR are 0.99 and 0.95 respectively.

We use $\beta = \sigma$ (σ is the standard deviation of the noise) and the Langrange multiplier $\lambda = 0.005$ [46]. Figures 4.7.1(e) and 4.7.1(f) represent the denoised signals after 80 iterations with SNR ≈ 1.1 and 1.12 respectively.

We have performed many other experiments on 1D signals obtaining similar results.



(a) (b) (c)

(d) (e) (f)

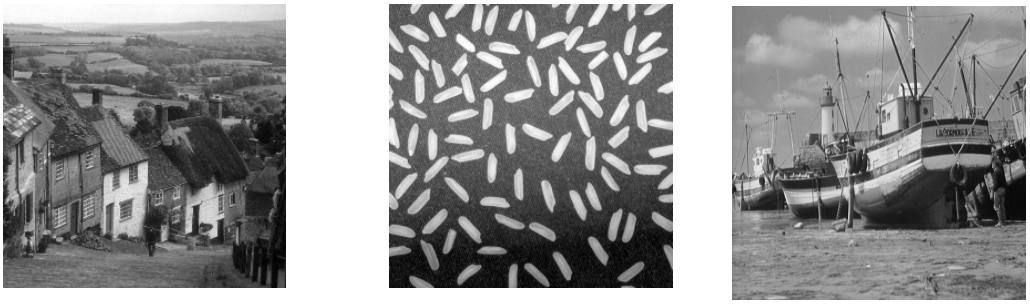
Figure 4.7.1: (a)-(b) Original signals; (c)-(d) corresponding noisy signals; (e)-(f) corresponding denoised signals.

4.8 Numerical experiments for time dependent model for 2D

In this Section, we have used three gray scale images, Goldhill (256×256), Rice (256×256) and Boat (512×512) shown in Figure 4.8.1 for our denoising experiments.

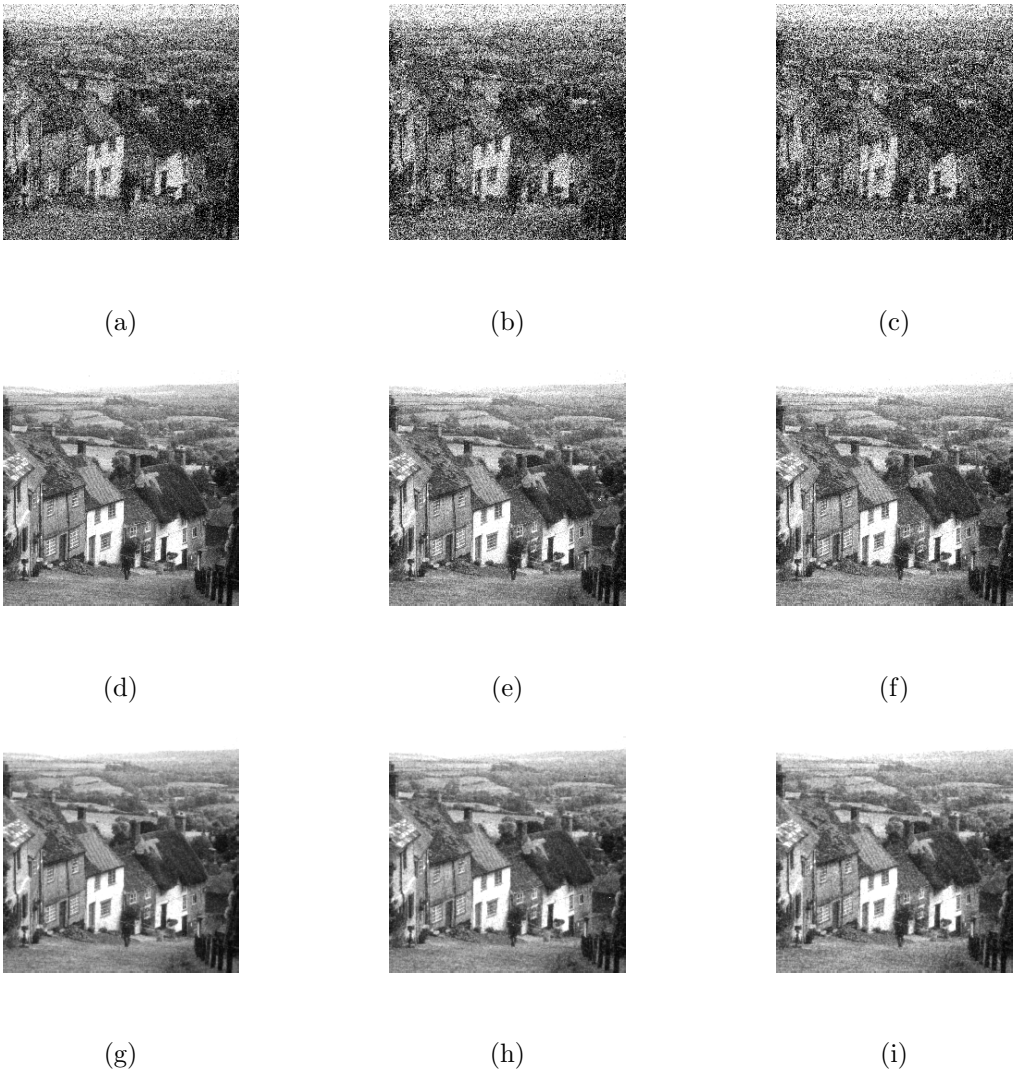
When Gaussian white noise with mean zero and variance σ^2 is added to the original images, we get noisy images. In our experiment, we have considered the images corrupted with different levels of Gaussian noise. Figure 4.8.2(a)-(c), Figure 4.8.3(a)-(c) and Figure 4.8.4(a)-(c) contain noisy images with different levels of Gaussian noise. The results obtained by using models (4.5.6) and (4.5.7) are shown in Figure 4.8.2-4.8.4 and Tables 1, 2 and 3. We have taken Lagrange multiplier $\lambda = 0.85$ as was used in references [15] and [17]. We can choose $\beta = 10^{-32}$ [17], the smallest positive machine number.

The values of PSNR obtained using model (4.5.7) given in Tables 1, 2 and 3 are larger than that of using model (4.5.6) at the same iteration number. Thus based on PSNR values and also on human perception, we conclude that the model (4.5.7) gives better denoised images than that of model (4.5.6).



(a) (b) (c)

Figure 4.8.1: Original Test Images used for different experiments (a) Goldhill: 256×256, (b) Rice: 256×256 and (c) Boat: 512×512.

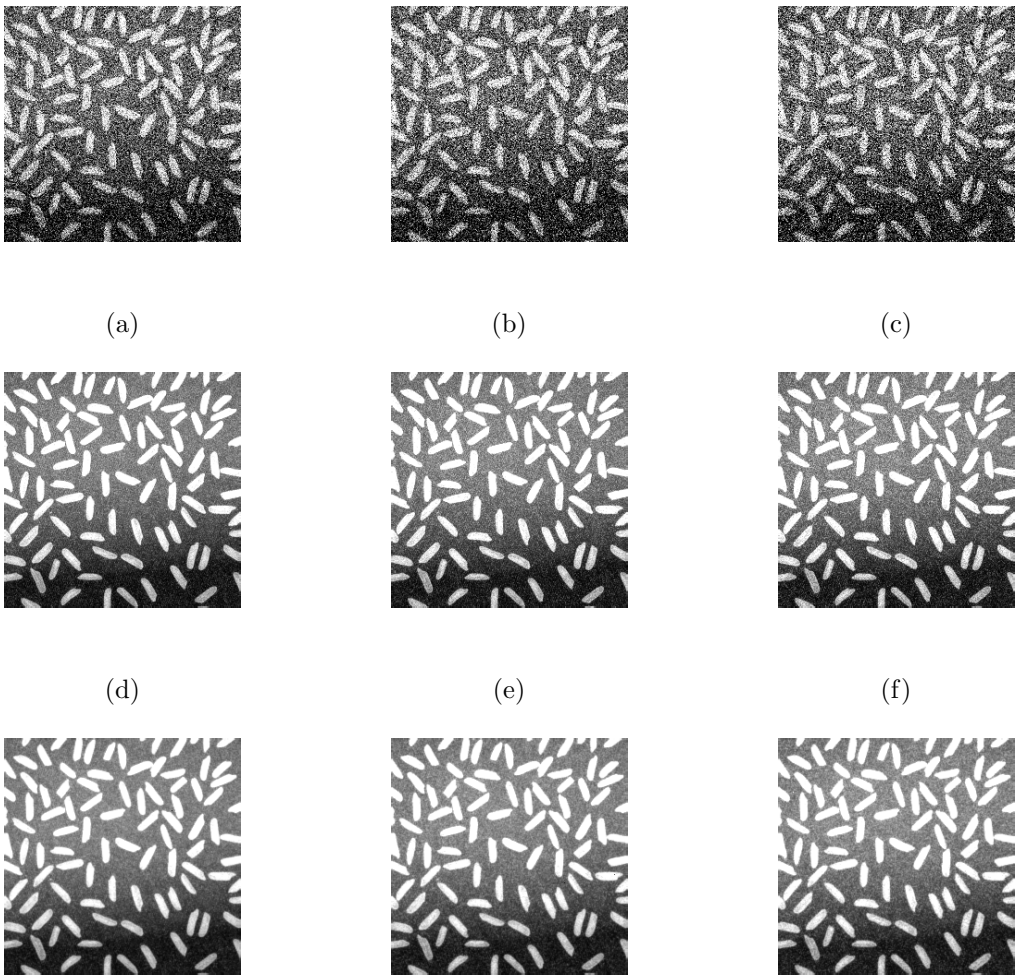


(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 4.8.2: (top row) Noisy Goldhill images with different levels of Gaussian noise (a)-(c), $\sigma^2 = 0.06,\ 0.08,\ 0.10$, respectively; (second row) (d)-(f) corresponding denoised images by model (4.5.6); (third row) (g)-(i) by model (4.5.7).

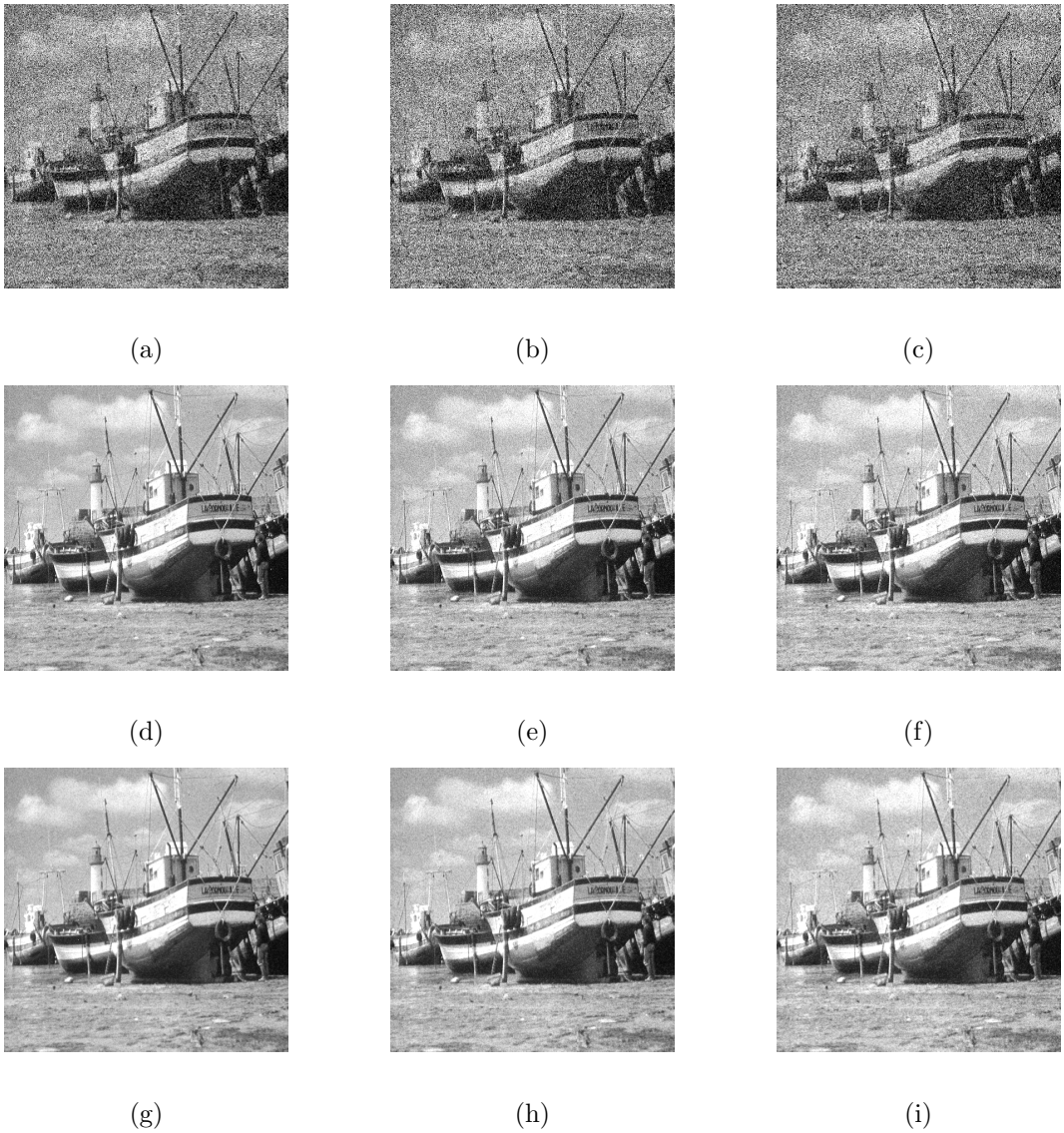


(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 4.8.3: (top row) Noisy Rice images with different levels of Gaussian noise (a)-(c), $\sigma^2 = 0.06,\ 0.08,\ 0.10$, respectively; (second row) (d)-(f) corresponding denoised images by model (4.5.6); (third row) (g)-(i) by model (4.5.7).



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 4.8.4: (top row) Noisy Boat images with different levels of Gaussian noise (a)-(c), $\sigma^2 = 0.06,\ 0.08,\ 0.10$, respectively; (second row) (d)-(f) corresponding denoised images by model (4.5.6); (third row) (g)-(i) by model (4.5.7).

Table 4.8.1: Results obtained by using models (4.5.6) and (4.5.7) applied to the images in Figure 4.8.2 with three different levels of Gaussian noise ($\sigma^2 = 0.06$, 0.08 and 0.10).

Images	PSNR (Noisy Images)	Images	PSNR (4.5.6)	Images	PSNR (4.5.7)
Fig. 4.8.2(a)	13.18	Fig. 4.8.2(d)	18.79	Fig. 4.8.2(g)	19.30
Fig. 4.8.2(b)	12.23	Fig. 4.8.2(e)	17.43	Fig. 4.8.2(h)	18.06
Fig. 4.8.2(c)	11.52	Fig. 4.8.2(f)	16.35	Fig. 4.8.2(i)	17.10
-	-	No. of iterations	5	No. of iterations	5

Table 4.8.2: Results obtained by using models (4.5.6) and (4.5.7) applied to the images in Figure 4.8.3 with three different levels of Gaussian noise ($\sigma^2 = 0.06$, 0.08 and 0.10).

Images	PSNR (Noisy Images)	Images	PSNR (4.5.6)	Images	PSNR (4.5.7)
Fig. 4.8.3(a)	13.36	Fig. 4.8.3(d)	19.10	Fig. 4.8.3(g)	19.39
Fig. 4.8.3(b)	12.38	Fig. 4.8.3(e)	17.85	Fig. 4.8.3(h)	18.26
Fig. 4.8.3(c)	11.64	Fig. 4.8.3(f)	16.86	Fig. 4.8.3(i)	17.38
-	-	No. of iterations	5	No. of iterations	5

Table 4.8.3: Results obtained by using models (4.5.6) and (4.5.7) applied to the images in Figure 4.8.4 with three different levels of Gaussian noise ($\sigma^2 = 0.06$, 0.08 and 0.10).

Images	PSNR (Noisy Images)	Images	PSNR (4.5.6)	Images	PSNR (4.5.7)
Fig. 4.8.4(a)	12.96	Fig. 4.8.4(d)	17.07	Fig. 4.8.4(g)	17.50
Fig. 4.8.4(b)	11.97	Fig. 4.8.4(e)	15.75	Fig. 4.8.4(h)	16.28
Fig. 4.8.4(c)	11.27	Fig. 4.8.4(f)	14.73	Fig. 4.8.4(i)	15.33
-	-	No. of iterations	5	No. of iterations	5

4.9 Conclusion

In this Chapter, we have presented a second order PDE based new nonlinear diffusion model (4.2.4) and a new time dependent model (4.5.7) to solve the nonlinear total variation problem for image denoising in 2D. The main idea is to apply a priori smoothness on the solution images. The model (4.2.4) gives larger PSNR values than that of model (4.2.3) even at relatively small iteration numbers and model (4.5.7) gives larger PSNR values than that of model (4.5.6), at the same iteration numbers. Besides, a new time dependent model (4.6.4) to solve the signal denoising in 1D has also been given.

Chapter 5

A time dependent model for image restoration with forward-backward diffusivities

5.1 Introduction

Image restoration is a fundamental problem in both image processing and computer vision with numerous application. Given a blurry and noisy image $u_0 : \Omega \to \mathbb{R}$,

$$u_0 = k * u + n, \tag{5.1.1}$$

where Ω is a bounded open set in $\mathbb{R}^2$, u_0 is the observed image, u is the original image, k is the point spread function (PSF) usually called blurred kernel and n is additive white noise assumed to be close to Gaussian. The values $n(i, j)$ of n at the pixels (i, j) are independent random variables, each with a Gaussian distribution of zero mean and variance σ^2.

Consider the degradation model (5.1.1), taking the Fourier transform we arrive at

$$\hat{u_0} = \hat{u} \; . \; \hat{k} + \hat{n}. \tag{5.1.2}$$

To recover u, we need to deconvolve, i.e., we have to divide (5.1.2) by $\hat{k}$ and then apply the inverse Fourier transform. This procedure is ill-posed. If k is smooth, high frequencies tend quickly to zero, implying that those frequencies in $\hat{u_0}$ get amplified, and the above model, in spite of its simplicity, is far from efficient.

The Total variation (TV) model is proposed by many researchers but the first

approach in this regard recalls the name Rudin et al. [61]. In 1992, they proposed a constrained optimization type numerical scheme for image denoising. The solution of the imposed problem is obtained by gradient projection method. In 1994, Rudin and Osher [60] proposed another model for image restoration. The researchers looked towards the idea of total variation and have given improved and fast versions of the TV technique [15, 17, 47, 60, 61, 76]. Total variation denoising is a popular method and is considered as a bottom line for edge preserving in image restoration. A large statement is that this method is able to restore sharp edges but at the same time, might met up with some staircasing (i.e., spurious edges) in plane regions. In spite of the fact that the variational problem is convex, the Euler-Lagrange equation is non linear and ill-conditioned. Linear semi-implicit fixed point procedures devised by Vogel and Oman [76], and interior-point primal dual implicit quadratic methods by Chain, Golub and Mulet [15], were introduced to solve the models. Deconvolution with total variation regularisation by variational approaches is studied in [7, 16, 47, 85] which deals with both classes of problems arising in non blind and blind deconvolution.

The rest of this Chapter is organized as follows: Section 5.2 reviews the major PDE based image deblurring and denoising models. Section 5.3 describes the choices of the diffusivity functions. In Section 5.4, we discuss our models and existing models applied on various types of gray scale images. To quantify results, the experimental values in terms of improvement in signal to noise ratio (ISNR) are given in Tables 5.4.1-5.4.4. Section 5.5 concludes the chapter.

5.2 Variational/ PDE-based deconvolution models

In this Section, we present the image restoration models [60, 61] and our new nonlinear diffusion models. The total variational functional introduced by Rudin et al. [61] and Rudin and Osher [60] is given as:

$$\begin{aligned} &\min_u \int_\Omega |\nabla u| \; dx \; dy, \\ &\text{Subject to } ||k * u - u_0||^2_{L^2} = |\Omega|\sigma^2. \end{aligned} \tag{5.2.1}$$

The equivalent unconstrained minimization problem of (5.2.1) can be written as:

$$\min_u \int_\Omega \left(|\nabla u| + \frac{\lambda}{2} \int_\Omega (k * u - u_0)^2 \right) dx\ dy. \tag{5.2.2}$$

The Euler-Lagrange equation is given by

$$0 = -\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda k * (k * u - u_0). \tag{5.2.3}$$

Since (5.2.3) is not well defined at points where $\nabla u = 0$, due to the presence of the term $\frac{1}{|\nabla u|}$, it is common to slightly perturb the TV algorithm to become

$$\int_\Omega |\nabla u|_\beta\ dx\ dy = \int_\Omega \sqrt{u_x^2 + u_y^2 + \beta}\ dx\ dy, \tag{5.2.4}$$

where β is a small positive parameter [17].

Usually, time dependent approximations to the ill-conditioned Euler-Lagrange equation (5.2.3) are inefficient. This is because a very small time step is required when a simple explicit scheme is used.

Rudin et al. [60] introduced a time dependent model for image restoration which is given by

$$u_t = \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) - \lambda k * (k * u - u_0), \tag{5.2.5}$$

with $u(x, y, 0)$ given as initial data (the original blurred and noisy image u_0 is taken as initial guess) and homogeneous Neumann boundary conditions $\frac{\partial u}{\partial \vec{n}} = 0$ on the boundary of the domain. The total variation norm does not penalize discontinuities in u, and thus allows us to recover the edges of the original image. To overcome the computational difficulties of total variation (TV) restoration problems, Vogel et al. [76] devised linear semi implicit fixed point procedures and Chan et al. [15] gave a primal-dual implicit quadratic methods. These methods give good results when treating pure denoising problems, but methods become very ill-conditioned for the deblurring and denoising case where the computational cost is very high and parameter dependent. As t increases, we approach a restored version image, and the effect of the evolution should be edge enhancement and smoothing at small scales in order to remove the noise. This solution procedure is a parabolic equation with time as an evolution parameter and

resembles the gradient-projection method of Rosen [59].

Sometimes the model (5.2.5) converges very slowly to its steady state since the parabolic term is singular for small gradients. At the same time, the Courant-Friedrichs-Lewy (CFL) restriction for keeping stability must be noticed. In order to regularize the parabolic term we multiply the whole Euler-Lagrange equation (5.2.3) by the magnitude of the gradient and time evolution model (see the reference [47]) is given by

$$u_t = |\nabla u| \nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) - |\nabla u| \lambda k * (k * u - u_0), \tag{5.2.6}$$

with $u(x, y, 0)$ given as initial data (the original blurred and noisy image u_0 is taken as initial guess) and homogeneous Neumann boundary conditions $\frac{\partial u}{\partial \vec{n}} = 0$.

Applying a priori smoothness on the solution image, a time dependent model becomes,

$$u_t = |\nabla G_\sigma * u| \nabla . \left(\frac{\nabla G_\sigma * u}{|\nabla G_\sigma * u|} \right) - |\nabla G_\sigma * u| \lambda k * (k * G_\sigma * u - u_0), \tag{5.2.7}$$

with $u(x, y, 0)$ given as initial data and homogeneous Neumann boundary conditions as above. It should be noticed that (5.2.7) only replaces u in (5.2.6) by its estimate $G_\sigma * u$.

Witkin [83] noticed that the convolution of the signal with Gaussian at each scale was equivalent to solving the heat equation with the signal as initial datum. The term $(G_\sigma * \nabla u)(x, y, t) = (\nabla G_\sigma * u)(x, y, t)$, which appears inside the divergence term of (5.2.7), is simply the gradient of the solution at time σ of the heat equation with $u(x, y, 0)$ as initial datum. Then the restoration analysis associated with u_0 consists in solving the problem

$$\partial u(x, y, t)/\partial t = \triangle u(x, y, t), \qquad u(x, y, 0) = u_0(x, y).$$

The solution of this equation at time t is given by

$$u(x, y, t) = G_\sigma * u_0,$$

where G_σ is the Gaussian function.

In order to preserve the notion of scale in the gradient estimate, it is convenient

that this kernel G_σ depends on a scale parameter [45]. In fact, the function G_σ can be considered as "low-pass filter" or any smoothing kernel, i.e., a denoising technique is used before solving the nonlinear diffusion problem [3, 13].

The use of nonconvex regularization functionals in image restoration has been investigated by Welk et al. [82]. In general, variational deblurring and denoising of an image can be achieved by minimizing the energy functional

$$E(u) = \int_\Omega \psi(|\nabla u|^2)\ dx\ dy + \frac{\lambda}{2}\int_\Omega (k * u - u_0)^2\ dx\ dy. \tag{5.2.8}$$

The first integral is the smoothness term or regulariser and the second integral, the data term is the squared error of the reconstruction of the blurred image from the deblurred image. This data term arises naturally in the deconvolution context and is also used in the variational blind models in [7, 85].

The Euler-Lagrange equation associated with (5.2.8) with homogeneous Neumann boundary conditions, given by

$$\begin{cases} 0 = -\ \mathrm{div}(\psi^{'}(|\nabla u|^2)\nabla u) + \lambda\ k * (k * u - u_0), \quad x, y \in \Omega, \\ \\ \frac{\partial u}{\partial \vec{n}} = 0, \quad x, y \in \partial\Omega, \end{cases} \tag{5.2.9}$$

where $\partial\Omega$ is the boundary of Ω and $\vec{n}$ is the outward normal to $\partial\Omega$.

A gradient descent leading for $t \to \infty$ to a minimizer of E is given by

$$u_t = \mathrm{div}(g(|\nabla u|^2)\nabla u) - \lambda\ k * (k * u - u_0), \tag{5.2.10}$$

with homogeneous Neumann boundary conditions. It is also known as diffusion-reaction equation where the diffusion term with diffusivity $g(s^2) = \psi'(s^2)$ is related to the regulariser in the energy functional.

The model (5.2.10) converges very slowly to its steady state in explicit schemes for image restoration, for details we refer [46]. So we multiply the whole Euler-Lagrange equation (5.2.9) by the magnitude of the gradient and our nonlinear anisotropic diffusion model takes the following form:

$$u_t = |\nabla u|\mathrm{div}(g(|\nabla u|^2)\nabla u) - |\nabla u|\lambda\ k * (k * u - u_0). \tag{5.2.11}$$

Applying a priori smoothness on the solution image, our nonlinear anisotropic diffusion model becomes,

$$u_t = |\nabla G_\sigma * u|\text{div}(g(|\nabla G_\sigma * u|^2)\nabla G_\sigma * u) - |\nabla G_\sigma * u|\lambda\ k*(k*G_\sigma * u - u_0). \quad (5.2.12)$$

It should be noticed that (5.2.12) only replaces u in (5.2.11) by its estimate $G_\sigma * u$.

We still write $G_\sigma * u$ as u. Let u_{ij}^n be the approximation to the value $u(x_i, y_j, t_n)$, where

$$x_i = i\Delta x,\ y_j = j\Delta x,\ i, j = 1, 2, ..., N,$$

$$t_n = n\Delta t,\quad n \geq 1,$$

where $\Delta x, \Delta y$ and Δt are the spatial step sizes and the time step size respectively.

The explicit partial derivatives of models (5.2.6) and (5.2.7) can be expressed as:

$$u_t = \frac{u_{xx}(u_y^2 + \beta) - 2u_{xy}u_x u_y + u_{yy}(u_x^2 + \beta)}{(u_x^2 + u_y^2 + \beta)} - \sqrt{u_x^2 + u_y^2 + \beta}\ \lambda k*(k*u - u_0). \quad (5.2.13)$$

We define the derivative terms as,

$$u_{ij}^x = \frac{u_{i+1,j}^n - u_{i-1,j}^n}{2\Delta x}; \qquad u_{ij}^y = \frac{u_{i,j+1}^n - u_{i,j-1}^n}{2\Delta x};$$

$$u_{ij}^{xx} = \frac{u_{i+1,j}^n - 2u_{i,j}^n + u_{i-1,j}^n}{\Delta x^2}; \qquad u_{ij}^{yy} = \frac{u_{i,j+1}^n - 2u_{i,j}^n + u_{i,j-1}^n}{\Delta x^2};$$

$$u_{ij}^{xy} = \frac{u_{i+1,j+1}^n - u_{i-1,j+1}^n - u_{i+1,j-1}^n + u_{i-1,j-1}^n}{4\Delta x\Delta x}; \qquad u_{ij}^t = \frac{u_{i,j}^{n+1} - u_{i,j}^n}{\Delta t}.$$

We let,

$$r_{ij}^n = u_{ij}^{xx}((u_{ij}^y)^2 + \beta) - 2u_{ij}^{xy}u_{ij}^x u_{ij}^y + u_{ij}^{yy}((u_{ij}^x)^2 + \beta), \quad (5.2.14)$$

and

$$p_{ij}^n = ((u_{ij}^x)^2 + (u_{ij}^y)^2 + \beta). \quad (5.2.15)$$

Then (5.2.13) reads as follows:

$$u_{ij}^t = \frac{r_{ij}^n}{p_{ij}^n} - \sqrt{((u_{ij}^x)^2 + (u_{ij}^y)^2 + \beta)}\ \lambda\ k*(k*u_{ij}^n - u_{ij}^0), \quad (5.2.16)$$

with homogeneous Neumann boundary conditions.

The explicit partial derivatives of models (5.2.11) and (5.2.12) can be expressed as:

$$\begin{aligned} u_{ij}^t &= \frac{1}{2\Delta x}\, s_{ij}^n\, ((g_{i+1,j}^n + g_{i,j}^n)(u_{i+1,j}^n - u_{i,j}^n) - (g_{i,j}^n + g_{i-1,j}^n)(u_{i,j}^n - u_{i-1,j}^n)) \\ &\quad + \frac{1}{2\Delta x}\, s_{ij}^n\, ((g_{i,j+1}^n + g_{i,j}^n)(u_{i,j+1}^n - u_{i,j}^n) - (g_{i,j}^n + g_{i,j-1}^n)(u_{i,j}^n - u_{i,j-1}^n)) \\ &\quad - s_{ij}^n\, \lambda k * (k * u_{ij}^n - u_{ij}^0), \end{aligned} \tag{5.2.17}$$

where the diffusivity $g(|\nabla u|^2)$ is discretised by,

$$g_{ij}^n = \psi^{'}\left(\left(\frac{u_{i+1,j}^n - u_{i-1,j}^n}{\Delta x}\right)^2 + \left(\frac{u_{i,j+1}^n - u_{i,j-1}^n}{\Delta x}\right)^2\right),$$

and $s_{ij}^n = \sqrt{(u_{ij}^x)^2 + (u_{ij}^y)^2}$,
with homogeneous Neumann boundary conditions.

The explicit method is stable and convergent for $\frac{\Delta t}{\Delta x^2} \leq 0.5$, see [41].

5.3 Choice of the diffusivity

In the deconvolution process, the choice of the diffusivity g is very important. Take the simplest case, the constant diffusivity $g(s^2) = 1$. It gives an over-smoothed deblurring result because high gradients at edges of the reconstructed image are penalised over-proportionally. Moreover, in this case the whole deconvolution method is again linear and suffers from the artifacts.

Total variation (TV) diffusivity $g(s^2) = \frac{1}{|s|}$, in its regularised form $g(s^2) = \frac{1}{\sqrt{s^2+\epsilon^2}}$, is a popular choice, see for example references [7, 16, 47]. It enforces piecewise constant results and therefore encourages sharp edges in the image.

The Perona-Malik (PM) diffusivity $g(s^2) = (1 + \frac{s^2}{\gamma^2})^{-1}$, that is related to the non-convex regulariser $\psi(s^2) = \gamma^2 log(1 + \frac{s^2}{\gamma^2})$, where γ is a contrast parameter determines which steepness edges are enhanced in the gradient descent process, see the references [56, 78]. We have included Total variation and Perona-Malik diffusivity in our experiments.

5.4 Numerical implementation

In this Section, we perform numerical experiments on the gray scale images, Lena and Boat of (256×256) pixels. We first scale the intensities of the images into the range between zero and one before we begin our experiments. The Gaussian white noise is added by the normal imnoise function imnoise (I,'Gaussian', M, σ^2), for the mean M and variance σ^2, in Matlab. In our test, we will use the blurred signal to noise ratio (BSNR) is used to measure the ratio of the level of blurred kernel and the level of noise,

$$\text{BSNR} = 10 \log_{10} \left(\frac{\text{blurred signal variance}}{\text{noise variance}} \right) dB. \tag{5.4.1}$$

Improvement in the signal quality (ISNR) is used to measure the goodness of restored image:

$$\text{ISNR} = 10 \log_{10} \left(\frac{\sum_{i,j}^{n} [u_{ij} - (u_0)_{ij}]^2}{\sum_{i,j}^{n} [u_{i,j} - (u_{new})_{i,j}]^2} \right) dB, \tag{5.4.2}$$

where u_{new} is the restored image. That is, the value of ISNR is larger, the restored image is better.

Here we use Gaussian kernel, defined as

$$k_\alpha(x, y) = \frac{1}{2\pi\alpha^2} \exp^{\frac{-(x^2+y^2)}{2\alpha^2}}. \tag{5.4.3}$$

The size k_α of blurring operator is 5 and 11. We use the blurring parameter α=2 and 3 in the blurred kernel (5.4.3). We have taken Lagrange multiplier $\lambda = 0.85$, see references [15, 17]. We can choose $\beta = 10^{-32}$, the smallest positive machine number as set in reference [17]. We have used γ=5 in PM diffusivity and ϵ=0.01 in TV diffusivity for our experiments. For models (5.2.6) and (5.2.7), we have considered $\Delta t/\Delta x^2 = 0.01$ and have taken $\Delta t/\Delta x^2 = 0.1$ for models (5.2.11) and (5.2.12).

Figures 5.4.1(b) and 5.4.2(b) represent the blurred and noisy images of Lena and Boat respectively with the size of Gaussian blurred kernel 5, the blurring parameter 2 and additive Gaussian white noise $\sigma^2 = 0.002$ wherein their BSNR $\approx$ 13.35 and 12.19 respectively. Figures 5.4.3(b) and 5.4.4(b) represent the blurred and noisy images of Lena and Boat respectively with the size of Gaussian blurred kernel 11, the blurring parameter 3 and additive Gaussian white noise $\sigma^2 = 0.002$ wherein their BSNR $\approx$ 12.76 and 11.67 respectively.

(a) (b) (c)

(d) (e) (f)

(g) (h)

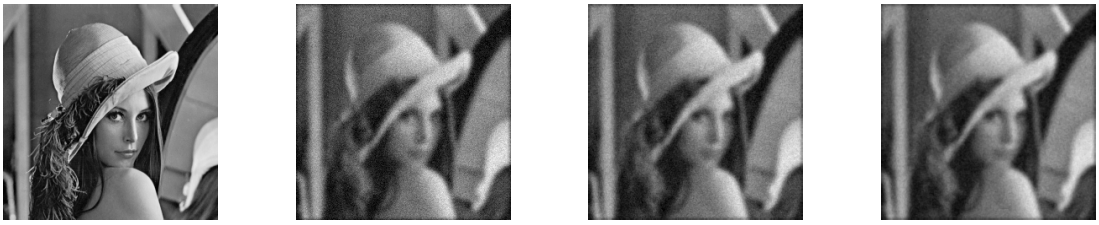
Figure 5.4.1: (a) Original Lena image, (b) blurred and noisy image with the size of Gaussian blur operator 5, blurring parameter 2 and Gaussian noise $\sigma^2 = 0.002$, (c) and (d) restored image by models (5.2.6) and (5.2.7) respectively, (e) and (f) restored image by models (5.2.11) and (5.2.12) respectively with PM diffusivity, (g) and (h) restored image by models (5.2.11) and (5.2.12) respectively with TV diffusivity.

(a) (b) (c)

(d) (e) (f)

(g) (h)

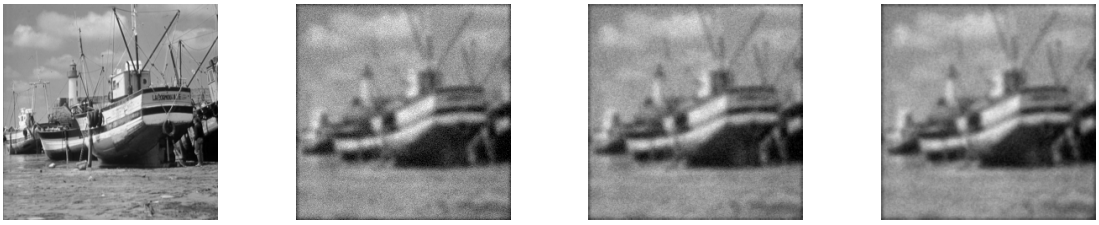
Figure 5.4.2: (a) Original Boat image, (b) blurred and noisy image with the size of Gaussian blur operator 5, blurring parameter 2 and Gaussian noise $\sigma^2 = 0.002$, (c) and (d) restored image by models (5.2.6) and (5.2.7) respectively, (e) and (f) restored image by models (5.2.11) and (5.2.12) respectively with Perona-Malik diffusivity, (g) and (h) restored image by models (5.2.11) and (5.2.12) respectively with TV diffusivity.



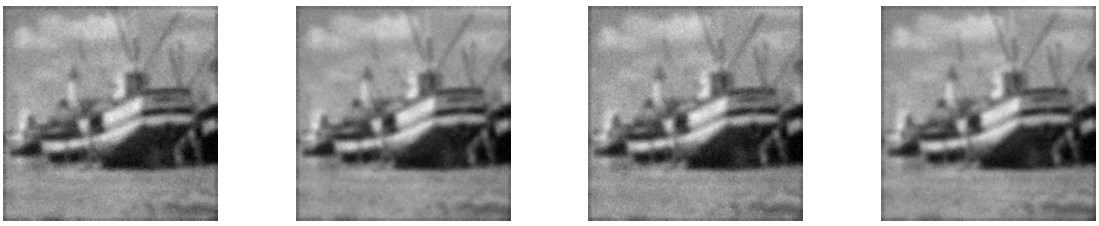
(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 5.4.3: (a) Original Lena image, (b) blurred and noisy image with $k_\alpha = 11$, $\alpha = 3$ and $\sigma^2 = 0.002$, (c) and (d) restored image by models (5.2.6) and (5.2.7) respectively, (e) and (f) restored image by models (5.2.11) and (5.2.12) respectively with PM diffusivity, (g) and (h) restored image by models (5.2.11) and (5.2.12) respectively with TV diffusivity.



(a) (b) (c) (d)



(e) (f) (g) (h)

Figure 5.4.4: (a) Original Boat image, (b) blurred and noisy image with $k_\alpha = 11$, $\alpha = 3$ and $\sigma^2 = 0.002$, (c) and (d) restored image by models (5.2.6) and (5.2.7) respectively, (e) and (f) restored image by models (5.2.11) and (5.2.12) respectively with PM diffusivity, (g) and (h) restored image by models (5.2.11) and (5.2.12) respectively with TV diffusivity.

Table 5.4.1: Results obtained by using models (5.2.6), (5.2.7), (5.2.11) and (5.2.12) applied to the image in Figure 5.4.1(b) with the size of Gaussian blur operator 5, blurring parameter 2 and Gaussian noise $\sigma^2 = 0.002$.

Images	ISNR	Images	ISNR (5.2.11)	Images	ISNR (5.2.12)
Fig. 5.4.1(c)	1.3768 (Model-5.2.6)	Fig. 5.4.1(e)	1.9702	Fig. 5.4.1(f)	2.2441
Fig. 5.4.1(d)	1.7361 (Model-5.2.7)	Fig. 5.4.1(g)	1.7714	Fig. 5.4.1(h)	1.9961
No. of iterations	25	No. of iterations	1000	No. of iterations	1000

Table 5.4.2: Results obtained by using models (5.2.6), (5.2.7), (5.2.11) and (5.2.12) applied to the image in Figure 5.4.2(b) with the size of Gaussian blur operator 5, blurring parameter 2 and Gaussian noise $\sigma^2 = 0.002$.

Images	ISNR	Images	ISNR (5.2.11)	Images	ISNR (5.2.12)
Fig. 5.4.2(c)	1.3528 (Model-5.2.6)	Fig. 5.4.2(e)	1.8835	Fig. 5.4.2(f)	2.1209
Fig. 5.4.2(d)	1.6863 (Model-5.2.7)	Fig. 5.4.2(g)	1.7354	Fig. 5.4.2(h)	1.9937
No. of iterations	25	No. of iterations	1000	No. of iterations	1000

Table 5.4.3: Results obtained by using models (5.2.6), (5.2.7), (5.2.11) and (5.2.12) applied to the image in Figure 5.4.3(b) with the size of Gaussian blur operator 11, blurring parameter 3 and Gaussian noise $\sigma^2 = 0.002$.

Images	ISNR	Images	ISNR (5.2.11)	Images	ISNR (5.2.12)
Fig. 5.4.3(c)	0.7447 (Model-5.2.6)	Fig. 5.4.3(e)	1.1221	Fig. 5.4.3(f)	1.2853
Fig. 5.4.3(d)	0.9608 (Model-5.2.7)	Fig. 5.4.3(g)	1.0284	Fig. 5.4.3(h	1.1943
No. of iterations	25	No. of iterations	1000	No. of iterations	1000

Table 5.4.4: Results obtained by using models (5.2.6), (5.2.7), (5.2.11) and (5.2.12) applied to the image in Figure 5.4.4(b) with the size of Gaussian blur operator 11, blurring parameter 3 and Gaussian noise $\sigma^2 = 0.002$.

Images	ISNR	Images	ISNR (5.2.11)	Images	ISNR (5.2.12)
Fig. 5.4.4(c)	0.7825 (Model-5.2.6)	Fig. 5.4.4(e)	1.2408	Fig. 5.4.4(f)	1.4155
Fig. 5.4.4(d)	1.0084 (Model-5.2.7)	Fig. 5.4.4(g)	1.1671	Fig. 5.4.4(h)	1.3585
No. of iterations	25	No. of iterations	1000	No. of iterations	1000

5.5 Conclusion

We have presented a new nonlinear anisotropic diffusion models for image restoration. The denoising technique is used to every solution image before solving the nonlinear diffusion problem. Our models (5.2.11) and (5.2.12) give better restoration results in comparison with other existing models (5.2.6) and(5.2.7). Nonlinear explicit schemes are used to discretize models (5.2.6), (5.2.7), (5.2.11) and(5.2.12). The models (5.2.11) and (5.2.12) give larger ISNR values than that of models (5.2.6) and (5.2.7), at different iterative numbers.

Chapter 6

Denoising method based on wavelet coefficients via diffusion equation

6.1 Introduction

In this Chapter, our aim is to reconstruct the wavelet coefficients using nonlinear diffusion reaction equation for image denoising. Interest in digital image processing methods stems from two principal application areas:

(a) Improvement of pictorial information for human interpretation.

(b) Processing of image data for storage, transmission, and representation for autonomous machine perception.

The crucial point in these approaches is to distinguish between important features that should be kept or even enhanced. Mathematically different ways have been used to model how a smooth image should look like, i.e., a certain smoothness can be formalised in terms of differentiability orders and small modulus of derivatives. This idea leads to regularisation methods and related partial differential equations. Another kind of smoothness assumption is to discard certain wavelet coefficients, leading to the popular wavelet shrinkage method.

For denoising case, the underlying model for the image degradation is an additive noise model. Thus in discrete setting, we have three images $u_0, u, n \in \mathbb{R}^2$ such that

$$u_0(x, y) = u(x, y) + n(x, y),$$

where $u_0(x, y)$ denote the observed noisy image, $u(x, y)$ is an known ideal image for $x, y \in \Omega$, Ω is a bounded open subset of $\mathbb{R}^2$ and $n(x, y)$ is additive white noise assumed

to be close to Gaussian. The values $n(i,j)$ of n at the pixels (i,j) are independent random variables, each with a Gaussian distribution of zero mean and variance σ^2.

The choice of the diffusivity c is very important in controlling the smoothing and even enhancement of edges. The Charbonnier diffusivity $g(s) = \frac{1}{\sqrt{1+(|s|^2/K^2)}}$, that is related to the convex regularizer $\psi(s^2) = \sqrt{K^4 + K^2 s^2} - K^2$, see references [18, 79], is used in our experiments.

Definition 6.1.1. *Let $m \in \mathbb{N}$. Then for $x \in \mathbb{R}$, a function $\psi(x)$ is called a mother wavelet of order m if it satisfies the following:*

(1) If $m = 0$, $\psi(x) \in L^\infty(\mathbb{R})$. If $m \geq 1$, then $\psi(x)$ and all its derivatives up to order m belongs to $L^\infty(\mathbb{R})$.

(2) $\psi(x)$ and all its derivatives up to order m decrease rapidly as $x \to \pm\infty$.

(3) For each $k \in \{0, \ldots, m\}$,

$$\int_{-\infty}^{\infty} x^k \psi(x) dx = 0.$$

(4) The collection $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ forms an orthonormal basis of $L^2(\mathbb{R})$, the $\psi_{j,k}$ being constructed from mother wavelet using $\psi_{j,k}(x) = 2^{\frac{j}{2}}\psi(2^j x - k)$.

6.2 Discrete wavelet transform

The discrete wavelet transform (DWT), proposed by Stefane Mallat in 1989, is an efficient algorithm for calculating the wavelet coefficients of a discrete series. The idea is to filter the series, using the high and low pass filters associated with the wavelet basis to obtain the wavelet coefficients. DWT can be treated in terms of a low pass filter $\mathcal{H} = \{h_k\}$ and a high pass filter $\mathcal{G} = \{g_k\}$, where the h_k and g_k are the coefficients of the filters. Consider a function u observed at $N = 2^J$ equally spaced time points $\{t_i = 0, \ldots, N-1\}$. Let $c_{J,i} = u(t_i)$ for $i = 0, \ldots, N-1$. The discrete wavelet transform of the series can be found by the relations:

$$c_{j-1,i} = \sum_n h_{n-2i} c_{j,n} \tag{6.2.1}$$

$$d_{j-1,i} = \sum_n g_{n-2i} c_{j,n}. \tag{6.2.2}$$

The resulting wavelet transform is the collection of the detail coefficients at each level together with the smooth coefficients at the zero level. This is a orthogonally transformed representation of the original series of length N. The $2i$ terms in the equations (6.2.1) and (6.2.2) is an alternative way of representing the decimation step of the DWT.
It is equivalent to filtering using $\mathcal{H}$ and $\mathcal{G}$ and then applying even dyadic decimation (selecting every even observation from the series).

Another way to design the DWT is to construct an orthogonal matrix W associated with the particular wavelet basis being used. The DWT can then be defined as the matrix multiplication of this orthogonal matrix with a vector of observation points, x. That is,

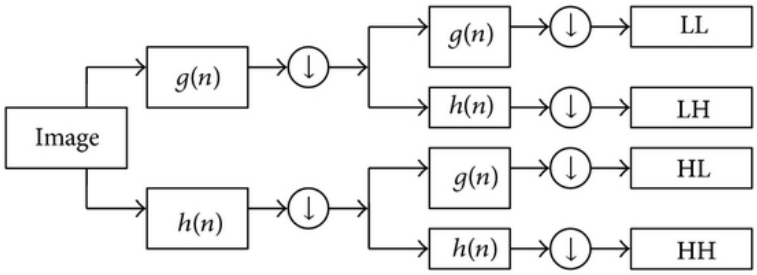
$$d = Wx$$

where d is a vector comprising of the both the discrete mother and father wavelet coefficients.

When the wavelet transform is found in terms of matrix multiplication, it is easy to look that the orthogonal matrix W can be inverted and so the inverse transform can be found. Computationally, the inverse transform takes the coarsest level coefficients and uses them to reconstruct the next finer level. Mathematically this can be written as

$$c_{j,k} = \sum_l h_{k-2l} c_{j-1,l} + \sum_l g_{k-2l} d_{j-1,l}.$$

The reconstruction can then be achieved by iterating this process and climbing the resolution levels back to the original data.



(a)

Figure 6.2.1: DWT structure of an image

6.3 Diffusion reaction equation based on wavelet coefficients

Originally the procedure to determine the minimum-norm solutions for ill-posed systems was given by Tikhonov [72], for regularization methods see the references [18, 54, 62, 63, 64]. Our exemplar of a variational method for denoising or simplifying an image u_0 can be achieved in [82] by minimizing the energy functional

$$E(u) = \int_{\Omega} \psi(|\nabla u|^2)\ dx\ dy + \frac{\lambda}{2} \int_{\Omega} (u - u_0)^2\ dx\ dy. \tag{6.3.1}$$

We see that it consists of two summands which are weighted by a constant regularization weight $\lambda \geq 0$. These two summands can be treated as data and smoothness terms respectively. To identify the minimizer of E we will follow the calculus of variations to obtain a necessary condition on the minimizer u, the so called Euler-Lagrange equation associated with (6.3.1) with homogeneous Neumann boundary conditions, given by

$$\begin{cases} 0 = -\text{div}(\psi^{'}(|\nabla u|^2)\nabla u) + \lambda\ (u - u_0) \quad x, y \in \Omega, \\ \\ \frac{\partial u}{\partial \vec{n}} = 0, \quad x, y \in \partial\Omega, \end{cases} \tag{6.3.2}$$

where $\partial\Omega$ is the boundary of Ω and $\vec{n}$ is the outward normal to $\partial\Omega$.

A gradient descent equation is given by

$$u_t = \text{div}(g(|\nabla u|^2)\nabla u) - \lambda\ (u - u_0), \tag{6.3.3}$$

with homogeneous Neumann boundary conditions. It is also known as diffusion-reaction equation where the diffusion term with diffusivity $g(s^2) = \psi'(s^2)$ is related to the regulariser in the energy functional.

Applying a priori smoothness on the solution image, our nonlinear diffusion model becomes,

$$u_t = \text{div}(g(|\nabla W u|^2)\nabla W u) - \lambda\ (W u - u_0), \tag{6.3.4}$$

where Wu is a wavelet coefficient. The process of removing noise by prior smoothness is achieved by means of wavelet decomposition and reconstruction.This does not harm the important features, for example edges of the image. We have proposed a new algorithm

for image denoising by prior smoothness and the design of algorithm is presented in Figure 2.

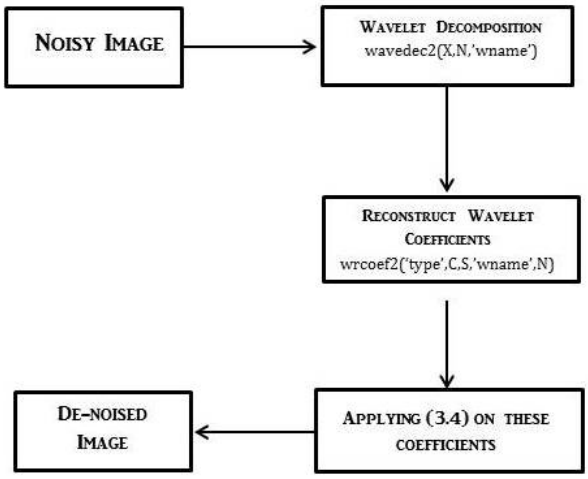


Figure 6.3.1: Algorithm for the proposed method

The process of decomposition is done by $X = wavedec2(u, `sym4', 1)$, where 'sym4' is a wavelet filter. After that, we reconstruct the wavelet coefficient by $Wu = wrcoef2(X, `sym4', 1)$ in matlab.

6.4 Discrete scheme

We still write Wu as u. Let u_{ij}^n be the approximation to the value $u(x_i, y_j, t_n)$, where

$$x_i = i\Delta x,\ y_j = j\Delta x,\ i, j = 1, 2, ..., N,$$

$$N\Delta x = 1,\quad t_n = n\Delta t,\quad n \geq 1,$$

where $\Delta x, \Delta y$ and Δt are the spatial step size and the time step size respectively.

The explicit partial derivatives of models (6.3.3) and (6.3.4) can be expressed as:

$$u_{ij}^t = \frac{1}{2\Delta x}\left((g_{i+1,j}^n + g_{i,j}^n)(u_{i+1,j}^n - u_{i,j}^n) - (g_{i,j}^n + g_{i-1,j}^n)(u_{i,j}^n - u_{i-1,j}^n)\right)$$

$$+\frac{1}{2\Delta x}\left((g_{i,j+1}^n + g_{i,j}^n)(u_{i,j+1}^n - u_{i,j}^n) - (g_{i,j}^n + g_{i,j-1}^n)(u_{i,j}^n - u_{i,j-1}^n)\right) - \ \lambda(u_{ij}^n - u_{ij}^0),$$

where the diffusivity $g(|\nabla u|)$ is discretised by,

$$g_{ij}^n = \psi' \left(\left(\frac{u_{i+1,j}^n - u_{i-1,j}^n}{\Delta x} \right)^2 + \left(\frac{u_{i,j+1}^n - u_{i,j-1}^n}{\Delta x} \right)^2 \right),$$

with homogeneous Neumann boundary conditions.

The explicit method is stable and convergent for $\frac{\Delta t}{\Delta x^2} \leq 0.5$, see [41].

6.5 Numerical implementation

We have used two gray scale images as shown in Figure 2.5.1. The pixel values of all images lie in interval [0, 255]. The Gaussian white noise is added by the normal imnoise function imnoise (I,'Gaussian', M, σ^2), i.e., the mean M and variance σ^2 in Matlab. We first scale the intensities of the images into the range between zero and one before we begin our experiments. We have taken $\Delta t/\Delta x^2 = 0.4$, Charbonnier diffusivity $K = 5$ and Lagrange multiplier $= 0.85$ as in [15] and [17], in our all experiments.

In our tests, we will use the peak signal to noise ratio (PSNR) as a criteria for the quality measurement of denoised images:

$$\text{PSNR} = 10\log_{10} \left(\frac{R^2}{\frac{1}{mn} \sum_{i,j}^{n} (u(i,j) - u_{new}(i,j))^2} \right), \tag{6.5.1}$$

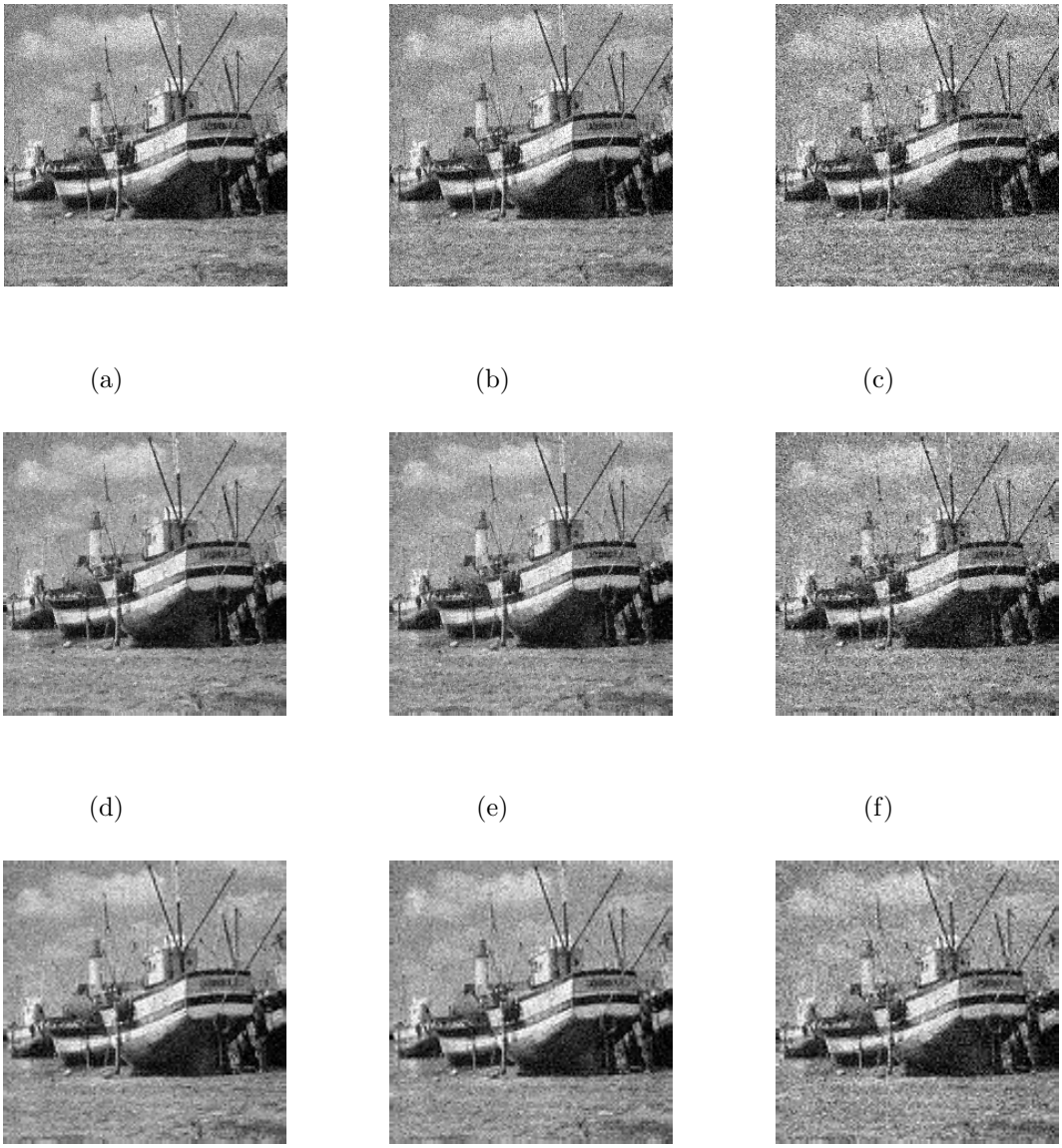
where $\{u(i,j) - u_{new}(i,j)\}$ is the difference of the pixel values between the denoised and original images.

(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 6.5.1: (top row) Noisy Lena images with different levels of Gaussian noise (a)-(c), $\sigma^2 = 0.008,\ 0.010,\ 0.020$, respectively; (second row) (d)-(f) corresponding denoised images by model (6.3.3); (third row) (g)-(i) by model (6.3.4).



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 6.5.2: (top row) Noisy Boat images with different levels of Gaussian noise (a)-(c), $\sigma^2 = 0.008,\ 0.010,\ 0.020$, respectively; (second row) (d)-(f) corresponding denoised images by model (6.3.3); (third row) (g)-(i) by model (6.3.4).

Table 6.5.1: Results obtained by using models (6.3.3) and (6.3.4) applied to the images in Figure 6.5.1 with three different levels of Gaussian noise (σ^2 = 0.008, 0.010 and 0.020).

Images	PSNR	Images	PSNR (6.3.3)	Images	PSNR (6.3.4)
Fig. 6.5.1(a)	21.18	Fig. 6.5.1(d)	24.35	Fig. 6.5.1(g)	25.34
Fig. 6.5.1(b)	20.22	Fig. 6.5.1(e)	23.28	Fig. 6.5.1(h)	24.80
Fig. 6.5.1(c)	17.44	Fig. 6.5.1(f)	19.86	Fig. 6.5.1(i)	22.57
-	-	No. of iterations	500	No. of iterations	50

Table 6.5.2: Results obtained by using models (6.3.3) and (6.3.4) applied to the images in Figure 6.5.2 with three different levels of Gaussian noise (σ^2 = 0.008, 0.010 and 0.020).

Images	PSNR	Images	PSNR (6.3.3)	Images	PSNR (6.3.4)
Fig. 6.5.2(a)	21.08	Fig. 6.5.2(d)	24.13	Fig. 6.5.2(g)	24.95
Fig. 6.5.2(b)	20.19	Fig. 6.5.2(e)	22.96	Fig. 6.5.2(h)	24.37
Fig. 6.5.2(c)	17.25	Fig. 6.5.2(f)	19.48	Fig. 6.5.2(i)	22.20
-	-	No. of iterations	500	No. of iterations	50

6.6 Conclusion

We have presented a new nonlinear anisotropic diffusion model via wavelet coefficients for image denoising. Here nonlinear explicit schemes are used to discretize models (6.3.3) and (6.3.4). The model (6.3.4) has achieved good results in comparison with the other process which are purely based on diffusion techniques. The fastness of our method is justified by the number of iterations.

Bibliography

[1] R.A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.

[2] L. Alvarez, J. Esclarn, *Image quantization using reaction-diffusion equations*, SIAM Journal on Applied Mathematics, 57(1), 153-175, 1997.

[3] L. Alvarez, P.L. Lions, J.M. Morel, *Image selective smoothing and edge detection by nonlinear diffusion II**, SIAM J. Numer. Anal., 29(3), 845-866, 1992.

[4] L. Ambrosia, N. Fusco, D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs, Oxford University Press, Oxford, UK, 2000.

[5] F. Andreu et al., C. Ballester, V. Caselles, J.M. Mazn, *Minimizing total variation flow*, Differential and Integral Equations, 14(3), 321-360, 2001.

[6] H. Andrew, B. Hunt, *Digital Image Restoration*, Prentice-Hall, Englewood Cliffs, NJ, 1977.

[7] L. Bar, N. Sochen, N. Kiryati, *Variational pairing of image segmentation and blind restoration*, In T. Pajdla, J. Matas, editors, Computer Vision ECCV, Lecture Notes in Computer Science, Springer Berlin, 3022, 166-177, 2004.

[8] T. Barbu, V. Barbu, V. Biga, D. Coca, *A PDE variational approach to image denoising and restoration*, Nonlinear Analysis: Real World Applications, 10(3), 1351-1361, 2009.

[9] C.A.Z. Barcelos, M. Boaventura, E.C. Silva, Jr., *A well-balanced flow equation for noise removal and edge detection*, IEEE Transactions on Image Processing, 12(7), 751-763, 2003.

[10] C.A.Z. Barcelos, Y. Chen, *Heat flows and related minimization problem in image restoration*, Computers and Mathematics with Applications, 39, 81-97, 2000.

[11] C. Blatter, *Wavelets a Primer*, Massachussetts: AK Peters, 1988.

[12] H. Brezis, *Analyse Fonctionnelle: Theorie et Applications*, Masson, Paris, 1987.

[13] F. Catté, P.L. Lions, J.M. Morel, T. Coll, *Image selective smoothing and edge detection by nonliear diffusion**, SIAM J. Numer Anal., 29, 182-193, 1992.

[14] A. Chambolle, P.L. Lions, *Image recovery via total variation minimization and related problems*, Numer. Math., 76, 167-188, 1997.

[15] T.F. Chan, G.H. Golub, P. Mulet, *A nonlinear primal-dual method for total variation-based image restoration*, SIAM J. Sci. Comput., 20(6), 1964-1977, 1999.

[16] T.F. Chan, C.K. Wong, *Total variation blind deconvolution*, IEEE Transactions on Image Processing, 7, 370-375, 1998.

[17] Q. Chang, I-Liang Chern, *Acceleration methods for total variation-based image denoising*, SIAM J. Sci. Comput., 25(3), 982-994, 2003.

[18] P. Charbonnier, L. Blanc-Feraud, G. Aubert, M. Barlaud, *Two deterministic half-quadratic regularization algorithms for computed imaging*, in Proceedings of IEEE International Conference on Image Processing, IEEE Computer Society Press, Austin, Tex, USA, 2, 168-172, November 1994.

[19] Y.G. Chen, Y. Giga, S. Goto, *Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations*, J. Differ. Geom., 33, 749-786, 1991.

[20] Y. Chen, B.C. Vemuri, Li Wang, *Image denoising and segmentation via nonlinear diffusion*, Computers and Mathematics with Applications, 39, 131-149, 2000.

[21] C.K. Chui, *An Introduction to Wavelets*, Boston: Academic Press, Inc., 1992.

[22] R.R. Coifman, D.L. Donoho, *Translation Invariant De-noising*, in Wavelets and Statistics, ed. A. Antoniadis, Springer-Verlag, Heidelberg, 1995.

[23] R. Crandall, *Project in Scientific Computation*, Springer-Verlag, New York, 197–198, 211-212, 1994.

[24] M.G. Crandall, H. Ishii, P.L. Lions, *User's guide to viscosity solution of second order partial differential equations*, Bull. Amer. math. Soc., 27, 1-67, 1992.

[25] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, 1992.

[26] L. Debnath, *Wavelet Transforms and Their Applications*, Birkhäuser, Boston, 2002.

[27] D.L. Donoho, *Denoising by soft-thresholding*, IEEE Trans. Inform. Theory, IT, 41, 613-627, 1995.

[28] L.C. Evans, *Partial Differential Equations*, Grad. stud. Math., AMS, 19, 1998.

[29] L.C. Evans, J. Spruck, *Motion of level sets by mean curvature*, I, J. Differ. Geom., 33, 635-681, 1991.

[30] H.Y. Gao, *Wavelet shrinkage denoised using the non-negative garrote*, Journal of Comput. and Graphical Statistics, 7(4), 469-488, 1998.

[31] H.Y. Gao, A.G. Bruce, *Wave shrink with firm shrinkage*, Statistica Sinica, 7, 855-874, 1997.

[32] Y. Giga, S. Goto, H. Ishii, M.H. Sato, *Comparison principle and convexity preserving properties for singular degenerate parabolic equation on unbounded domains*, Hokkaido University, preprint, 1990.

[33] R.C. Gonzalez, R.E. Woods, *Digital Image Processing*, Addison-Wesley Publishing Company, 1993.

[34] B.R. Hunt, *The application of constrained least squares estimation to image restoration by digital computer*, IEEE Trans. Comput., 22(9), 805-812, 1973.

[35] V. Hutson, J.S. Pym, *Applications of Functional Analysis and Operator Theory*, Academic Press, 1980.

[36] T. Iijima, *Basic theory on normalization of pattern (In case of typical one dimensional pattern)*, Bulletin of the Electrotechnical Laboratory, 26, 368-388, 1962.

[37] G. Kaiser, *A Friendly Guide to Wavelets*, Boston-Basel-Berlin: Birkhäuser, 1994.

[38] S.L. Keeling, R. Stollberger, *Nonlinear anisotropic diffusion filters for wide range edge sharpening*, Inverse Problem, 18, 175-190, 2002.

[39] J.J. Koenderink, *The Structure of Images*, Biological Cybernetics, 1984.

[40] O.A. Ladysheuskaya, V.A. Solomikov, N.N. Ural'tseva, *Linear and Quasi linear equation of parabolic type*, American Mathematical Society, Providence, RI 1968.

[41] L. Lapidus, G.F. Pinder, *Numerical solution of partial differential equations in science and engineering*, SIAM Rev., 25(4), 581-582, 1983.

[42] Jichun Li, *Finite element analysis for a nonlinear diffusion model in image processing*, Applied Mathematics Letters, 15, 197-202, 2002.

[43] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1998.

[44] S. Mallat, *A theory for multiresolution signal decomposition: the wavelet representation*, IEEE transactions on pattern analysis and machine intelligence, 11(7), 674-693, 1989.

[45] A. Marquina, *Inverse scale space methods for blind deconvolution*, UCLA CAM Report, 06-36, 2006.

[46] A. Marquina, S. Osher, *Explict algorithms for a new time dependent model based on level set motion for nonlinear deblurring and noise removal*, SIAM. J. Sci. Comput., 22(2), 387-405, 2000.

[47] A. Marquina, S. Osher, *A new time dependent model based on level set motion for nonlinear deblurring and noise removal*, In M. Nielsen, P. johansen, O. F. Olsen and J. Weickert, editors, scale-Space Theories in Computer Vision, Lecture notes in Computer Science, Springer Berlin, 1688, 429-434, 1999.

[48] Y. Meyer, *Wavelets, Algorithms and Applications*, Philadelphia: SIAM, 1993.

[49] Y. Meyer, *Wavelets and Operators*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 37, 1992.

[50] P. Mrázek, J. Wieckert, *From two dimensional nonlinear diffusion to coupled haar wavelet shrinkage*, J. Visual Comm. & Image representation, 18, 162-175, 2007.

[51] P. Mrázek, J. Wieckert, G. Steidl, *Correspondence between wavelet shrinkage and nonlinear diffusion**, L.D. Griffin and M. Lillholm (Eds.): Scale-Space, LNCS, Springer-Verlag Berlin Heidelberg, 2695, 101-116, 2003.

[52] M.K. Ng, R.H. Chan, W.C. Tang, *A fast algorithm for deblurring models with Neumann boundary conditions*, SIAM J. Sci. Comput., 21, 851-866, 1999.

[53] L. Nirenberg, *Topics in Nonlinear Functional Analysis*, Lecture notes, New York Univesity, New York, 1974.

[54] N. Nordstrom, *Biased anisotropic diffusion a unified regularization and diffusion approach to edge detection*, Image and Vision Computing, 8, 318-327, 1990.

[55] S. Osher, L.I. Rudin, *Feature oriented image enhancement using shock filters*, SIAM J. Num. Anal., 27(4), 919-940, 1990.

[56] P. Perona, J. Malik, *Scale space and edge detection using anisotropic diffusion*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 12, 629-639, 1990.

[57] D.L. Phillips, *A technique for the numerical solution of certain integral equations of the first kind*, Journal of the Association for Computing Machinery (ACM), 9(1), 84-97, 1962.

[58] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes in FORTRAN*, Cambridge University Press, New York, 498-499, 1992.

[59] J.G. Rosen, *The gradient projection method for nonlinear programming, Part II, nonlinear constraints*, J. Soc. Indust. Appl. Math., 9(4), 514-532, 1961.

[60] L. Rudin, S. Osher, *Total variation based image restoration with free local constraints*, Proc. IEEE Int. Conf. Imag. Proc., 31-35, 1994.

[61] L. Rudin, S. Osher, E. Fatemi, *Nonlinear total variation based noise removal algorithm*, Phys. D., 60, 259-268, 1992.

[62] O. Scherzer, *Denoising with higher order derivatives of bounded variation and an application to parameter estimation*, Computing, 60(1), 127, March 1998.

[63] O. Scherzer, J. Weickert, *Relations between regularization and diffusion filtering*, Journal of Mathematical Imaging and Vision, 12, 43-63, 2000.

[64] C. Schnörr, *Unique reconstruction of piecewise smooth images by minimizing strictly convex non-quadratic functionals*, Journal of Mathematical Imaging and Vision, 4, 189-198, 1994.

[65] S. Serra-Capizzano, *A note on antireflective boundary conditions and fast deblurring models*, SIAM J. Sci. Comput., 25, 1307-1325, 2003.

[66] Y. Shi, Q. Chang, *New time dependent model for image restoration*, Applied Mathematics and Computation, 179(1), 121-134, 2006.

[67] M. Soner, *Motion of a set by the curvature of its mean boundary*, Journal of differential equations, 101(2), 313-372, 1993.

[68] G. Strang, *On the construction and comparison of difference schemes*, SIAM J. numer. Anal., 5(3), 506-517, 1968.

[69] G. Strang, *Accurate partial difference methods II. Non linear problems*, Numer. Math., 6, 37-46, 1964.

[70] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, 1997.

[71] D. Strong, *Adaptive Total Variation Minimizing Image Restoration*, Ph.D. Thesis, UCLA Mathematics Department, USA, August 1997.

[72] A. N. Tikhonov, *Solution of incorrectly formulated problems and the regularization method*, Soviet Mathematics Doklady, 4(2), 1035-1038, 1963.

[73] S. Twomey, *On the numerical solution of Fredholm integral equations of the first kind by the inversion of the linear system produced by quadrature*, Journal of the Association for Computing Machinery (ACM), 10(1), 97-101, 1963.

[74] S. Twomey, *The application of numerical filtering to the solution of integral equations encountered in indirect sensing measurements*, J. Franklin Inst., 279, 95-109, 1965.

[75] M. Vetterli, C. Herley, *Wavelets and filter banks: Theory and design*, IEEE Transactions on Signal Processing, 40, 2207-2232, 1992.

[76] C.R. Vogel, M.E. Oman, *Iterative methods for total variation denoising*, SIAM j. Sci. Comput., 17, 227-238, 1996.

[77] David F. Walnut, *An Introduction to Wavelet Analysis*, Birkhauser, Boston, 2001.

[78] J. Weickert, *Anisotropic Diffusion in Image Processing*, Teubner, Stuttgart, 1998.

[79] J. Weickert, *A Review of nonlinear diffusion filtering*, In scale space LNCS. Springer Berlin, 1252, 3-28, 1997.

[80] J. Weickert, B.M.H. Romeny, *Efficient and reliable schemes for nonlinear diffusion filtering*, IEEE Transactions on Image Processing, 7(3), 398-410, 1998.

[81] M. Welk, G. Steidl, J. Weickert, *Locally analytic schemes: A link between diffusion filtering and wavelet shrinkage*, Appl. Comput. Harmon. Anal., 24, 195-224, 2008.

[82] M. Welk, D. Theis, T. Brox, J. Weickert, *PDE-based deconvolution with fordward-backward diffusivities and diffusion tensors*, In scale space, LNCS., Springer, Berlin, 585-597, 2005.

[83] A.P. Witkin, *Scale-space filtering*, Proc. IJCAI, Karlsruhe, 1019-1021, 1983.

[84] P. Wojtaszczyk, *A Mathematical Introduction to Wavelets*, London Mathematical Society, Students Texts 37, Cambridge University Press, 1997.

[85] Y.L. You, M. Kaveh, *Anisotropic blind image restoration*, In Proc. IEEE International Conference on Image Processing, Lausanne, Switzerland, 2, 461-464, Sept. 1996.

[86] Y.L. You, W. Xu, A. Tannenbaum, M. Kaveh, *Behavioral analysis of anisotropic diffusion in image processing*, IEEE trans. Image Process., 5, 1539-1553, 1996.

Research Publications

Paper-1

Journal of Signal and Information Processing, 2015, 6, 28-38
Published Online February 2015 in SciRes. http://www.scirp.org/journal/jsip
http://dx.doi.org/10.4236/jsip.2015.61003

A Time Dependent Model for Image Denoising

Santosh Kumar, Mohammad Kalimuddin Ahmad

Department of Mathematics Aligarh Muslim University, Aligarh, India
Email: skykumar87@gmail.com, ahmad_kalimuddin@yahoo.co.in

Received 27 January 2015; accepted 16 February 2015; published 25 February 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
http://creativecommons.org/licenses/by/4.0/
Open Access

Abstract

In this paper, we propose a new time dependent model for solving total variation (TV) minimization problem in image denoising. The main idea is to apply a priori smoothness on the solution image. This is a constrained optimization type of numerical algorithm for removing noise from images. The constraints are imposed using Lagrange's multipliers and the solution is obtained using the gradient projection method. 1D and 2D numerical experimental results by explicit numerical schemes are discussed.

Keywords

Total Variation, Image Denoising, Signal Denoising

1. Introduction

In many image processing problems, a denoising step is required to remove noise or spurious details from corrupted images. The presence of noise in images is unavoidable. It may be introduced at the stage of image formation like image recording, image transmission, etc. These random distortions make it difficult to perform any required image analysis. For example, the feature oriented enhancement introduced in [1] is very effective in restoring blurry images, but it can be "*frozen*" by an oscillatory noise component. Even a small amount of noise is harmful when high accuracy is required, especially in case of medical images.

In practice, to estimate a true signal in noise, the most frequently used methods are based on the least squares criteria. This procedure is L^2-norm dependent. L^2-norm based regularization is known to remove high frequency components in denoised images and make them appear smooth.

Most of the classical image deblurring or denoising techniques, due to linear and global approach, are

Paper-2

Gmail santosh kumar <skykumar87@gmail.com>

Re: Res paper for IJIAM

1 message

Dr Kalimuddin Ahmad <ahmad_kalimuddin@yahoo.co.in> Sun, Feb 14, 2016 at 6:14 PM
To: Rashmi Bhardwaj <rashmibha@gmail.com>
Cc: Santosh Kumar <skykumar87@gmail.com>, Abul Hasan <siddiqi.abulhasan@gmail.com>

Dear Prof Rashmi,

Thanks for your mail. Soon we will be sending you the tex file together with the corresponding files in journals format which I suppose is the format of springer journals.

With kind regards,

Kalimuddin Ahmad
Deptt of Maths, AMU

Sent from Yahoo Mail on Android

On Sun, Feb 14, 2016 at 5:31 PM, Rashmi Bhardwaj <rashmibha@gmail.com> wrote:

> Dear Dr Ahmad,
>
> On behalf of Indian Journal of Industrial and Applied Mathematics , we would like to confirm that your article titled "An efficient PDE-based model for image restoration" authored by Santosh Kumar and M.K. Ahmad has been accepted by the editorial board which will publish in the forthcoming issue of the Indian Journal of Industrial and Applied Mathematics.
>
> Kindly send the doc file of the paper for further processing.
>
> with best wishes
>
> **Prof. (Mrs.) Dr. Rashmi Bhardwaj**
> Professor & Incharge Mathematics Section,
> University School of Basic & Applied Sciences,
> B-504, Nonlinear Dynamics Research Lab,
> Guru Gobind Singh Indraprastha University,
> Sector 16 C, Dwarka, Delhi 110078, India.
> +91-9868818880, +91-9013534105
> 011-25302415
> rashmibha@gmail.com rashmib22@gmail.com

Noname manuscript No.
(will be inserted by the editor)

An efficient PDE-based model for image restoration

Santosh Kumar · M.K. Ahmad

Received: date / Accepted: date

Abstract In this paper, we propose new time dependent model for solving total variation (TV) minimization problem in image restoration. The main idea is to apply a priori smoothness on the solution image. The total variation of the image is minimized subject to constraints involving the point spread function (PSF) of the blurring process and the statistics of the noise. The blurring operator provides useful information in restoration. The constraints are implemented using Lagrange's multipliers. The solution is obtained using the gradient-projection method of Rosen [16]. We present proof of the existence, uniqueness and stability of the viscosity solution of our model. The results of our model using explicit numerical schemes are compared with other known image restoration models.

Keywords Total variation norm · image restoration · nonlinear diffusion
2010 AMS Subject Classification: 65M06, 76R50, 68U10, 35D40

1 Introduction

Image restoration is a fundamental problem in both image processing and computer vision with numerous applications. An image can be interpreted as a real function defined on Ω, a bounded and open domain of $\mathbb{R}^2$ (for simplicity we will assume Ω to be the square domain henceforth). Our problem is to restore an image which is contaminated with noise and blured in such a way that the process should recover the edges of the images. Formation of a blurred and noisy image is typically modeled as

$$u_0 = H[u] + n, \tag{1.1}$$

where u_0 is the observed image, H is the PSF (point spread function) and also known as blur kernel and n is additive noise. For all practical purposes, n is taken as Gaussian white noise, i.e. the values $n_{i,j}$ of n at the pixels (i,j) are independent random variables, each with a Gaussian distribution of zero mean and variance σ^2. The blur operator H can be described by a Fredholm first kind integral operator

$$H[u(x,y)] = \int_{\Omega} u(s,r)h(x,s,y,r)ds\ dr, \tag{1.2}$$

Santosh Kumar
Department of Mathematics
Aligarh Muslim University, Aligarh-202002, India
E-mail: skykumar87@gmail.com

M.K. Ahmad
Department of Mathematics
Aligarh Muslim University, Aligarh-202002, India
E-mail: ahmad_kalimuddin@yahoo.co.in

Paper-3

Gmail santosh kumar <skykumar87@gmail.com>

AEAM: Submission Confirmation for AEAM-D-15-00027R1

2 messages

International Journal of Advances in Engineering Sciences (AEAM) <em@editorialmanager.com> Mon, Feb 15, 2016 at 8:00 PM
Reply-To: "International Journal of Advances in Engineering Sciences (AEAM)" <rajeswari.sundaram@springer.com>
To: Santosh Kumar <skykumar87@gmail.com>

Ref.: Ms. No. AEAM-D-15-00027R1
A time dependent model for image restoration with forward - backward Diffusivities

Dear Mr. Kumar,

International Journal of Advances in Engineering Sciences and Applied Mathematics has received your revised submission.

You may check the status of your manuscript by logging onto Editorial Manager at http://aeam.edmgr.com/.

Kind regards,

Editorial Office
International Journal of Advances in Engineering Sciences and Applied Mathematics

---

santosh kumar <skykumar87@gmail.com> Wed, Feb 17, 2016 at 12:02 PM
To: ahmad_kalimuddin@yahoo.co.in

[Quoted text hidden]

AEAM-D-15-00027_R1 (6).pdf
2799K

International Journal of Advances in Engineering Sciences and Applied Mathematics

A time dependent model for image restoration with forward - backward Diffusivities

--Manuscript Draft--

Manuscript Number:	AEAM-D-15-00027R1
Full Title:	A time dependent model for image restoration with forward - backward Diffusivities
Article Type:	S.I. : PDE Models and Computation
Corresponding Author:	Santosh Kumar, M.Sc Aligarh Muslim University Aligarh, Uttar Pradesh INDIA
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Aligarh Muslim University
Corresponding Author's Secondary Institution:	
First Author:	Santosh Kumar, M.Sc
First Author Secondary Information:	
Order of Authors:	Santosh Kumar, M.Sc
	Md. Kalimuddin Ahmad, Ph.D.
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	In image processing community, image denoising and deblurring, data compression, edge detection etc. are the most challenging and basic hurdles. Inspired by A. Marquina et al. [9], Y. Shi et al. [15] and M. Welk et al. [19], we address new nonlinear anisotropic diffusion models for image denoising and deblurring. To judge our models, we compared with the existing models of A. Marquina et al. [9], Y. Shi et al. [15] and M. Welk et al. [19]. 2D numerical experimental results by explicit numerical schemes are discussed.
Response to Reviewers:	We have carefully gone through the comments of learned reviewers and incorporated / modified the manuscript as per their suggestions. I am thankful to reviewers for their fruitful suggestions that resulted into the improvement of our paper. Many typos are corrected and language of paper is improved. Para-wise reply: 1. We have rewritten the abstract. 2. We have added TV- model in introduction vis-a-vis de-noising / Page-2, Para-2. 3. TV model is added / Page-2, Para-1 of Section-2. 4. It is incorporated on Page-3 / lines 9-11. 5. Selection of parameters is explained on Page-7, Para-5 (last Para). 6. Yes, I do agree with the suggestions and that can be incorporated in our future work. 7. We shall take care of this in our future work. 8. Yes, we have checked and revised the paper as per reviewers comments.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

International Journal of Advances in Engineering Sciences and Applied Mathematics

Copyright Transfer and Financial Disclosure/Conflict of Interest Statement

The copyright to this article is transferred to the Indian Institute of Technology Madras (for U.S. government employees: to the extent transferable) effective if and when the article is accepted for publication. The author warrants that his/her contribution is original and that he/she has full power to make this grant. The author signs for and accepts responsibility for releasing this material on behalf of any and all co-authors. The copyright transfer covers the exclusive right to reproduce and distribute the article, including reprints, translations, photographic reproductions, microform, electronic form (offline, online) or any other reproductions of similar nature.

An author may self-archive an author-created version of his/her article on his/her own website and or in his/her institutional repository. He/she may also deposit this version on his/her funder's or funder's designated repository at the funder's request or as a result of a legal obligation, provided it is not made publicly available until 12 months after official publication. He/she may not use the publisher's PDF version, which is posted on www.springerlink.com, for the purpose of self-archiving or deposit. Furthermore, the author may only post his/her version provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The original publication is available at www.springerlink.com".

The author is requested to use the appropriate DOI for the article. Articles disseminated via www.springerlink.com are indexed, abstracted and referenced by many abstracting and information services, bibliographic networks, subscription agencies, library networks, and consortia. After submission of the agreement signed by the corresponding author, changes of authorship or in the order of the authors listed will not be accepted.

I, the undersigned corresponding author, also certify that I/we have no commercial associations (e.g., consultancies, stock ownership, equity interests, patent-licensing arrangements, etc.) that might pose a conflict of interest in connection with the submitted article, except as disclosed on a separate attachment. All funding sources supporting the work and all institutional or corporate affiliations of mine/ours are acknowledged in a footnote. *Please mention if a separate attachment is enclosed.*

Title of article:

A time dependent model for image restoration with forward-backward diffusivities.

Author(s):

Santosh Kumar and M. K. Ahmad

Author's signature:

Santosh Kumar

Date: 13.4.2015

Please sign this form and upload the scanned form at:
www.editorialmanager.com/aeam

Indian Institute of Technology Madras

http://www.springer.com/journal/12572/

Manuscript Click here to download Manuscript paper2.tex

Click here to view linked References

Noname manuscript No.
(will be inserted by the editor)

A time dependent model for image restoration with forward - backward Diffusivities

Santosh Kumar · M.K. Ahmad

Received: date / Accepted: date

Abstract In image processing community, image denoising and deblurring, data compression, edge detection etc. are the most challenging and basic hurdles. Inspired by A. Marquina et al. [9], Y. Shi et al. [15] and M. Welk et al. [19], we address new nonlinear anisotropic diffusion models for image denoising and deblurring. To judge our models, we compared with the existing models of A. Marquina et al. [9], Y. Shi et al. [15] and M. Welk et al. [19]. 2D numerical experimental results by explicit numerical schemes are discussed.

Keywords Total variation norm · image restoration · anisotropic diffusion

1 Introduction

Image restoration is a fundamental problem in both image processing and computer vision with numerous application. Given a blurry and noisy image $u_0 : \Omega \to \mathbb{R}$,

$$u_0 = k * u + n, \tag{1}$$

where Ω is a bounded open set in $\mathbb{R}^2$, u_0 is the observed image, u is the original image, k is the PSF (point spread function)-usually called blurred kernel and n is additive white noise assumed to be close to Gaussian. The values $n(i, j)$ of

S. Kumar
Department of Mathematics
Aligarh Muslim University
Aligarh, 202002, India
E-mail: skykumar87@gmail.com

M.K. Ahmad
Department of Mathematics
Aligarh Muslim University
Aligarh, 202002, India
E-mail: ahmad_kalimuddin@yahoo.co.in

Paper-4

Natural Sciences Publishing

Login | New user?

Applied and Computational Mathematics

An International Journal

ACM Home

ISSN: 1683-3511 (print version)

Editors in-Chief

Abbasov, Ali (Baku, Azerbaijan)
Aliev, Fikret (Baku, Azerbaijan)
Larin, Vladimir (Kiev, Ukraine)

Managing Editor
Gasimov, Yusif (Baku, Azerbaijan)

Journals Impact Factor is 0.452 (Thomson Reuter Agency Journal Citation Report - 2014).

http://acmij.az/view.php?lang=en&menu=0

Applied and Computational Mathematics an International Journal was founded in 2002 by the Ministry of Communications and Information Technology (Azerbaijan), Azerbaijan National Academy of Sciences and Institute of Applied Mathematics of Baku State University. This is a three-annual Journal that publishes high-quality peer-reviewed research papers on the board ranges of applied and computational mathematics.

Officialy cited as: ***Appl. Comput. Math.***

Impact Factor: **0.452 (2014)***
5-year Impact Factor: **0.592**
* Journal Citation Reports ®, Thomson Reuters

Source Normalized Impact per Paper (SNIP): **1.207**

ISSN: 1683-3511 (print version)

Copyright naturalspublishing.com. All Rights Reserved

Add Article

My Articles

1

Full Name	Send	Delete	Edit	Status
PDE-based nonlinear diffusion model for image denoising				

Copyright © naturalspublishing.com. All Rights Reserved

Appl. Comput. Math., V.xx, N.xx, 20xx, pp.xx-xx

PDE-BASED NONLINEAR DIFFUSION MODEL FOR IMAGE DENOISING

SANTOSH KUMAR[1] AND M.K. AHMAD[2]

ABSTRACT. In this paper, we propose a new model for image denoising. Second order partial differential equations have been studied as a useful tool for image denoising. Scale space and edge detection using anisotropic by Parona and Malik has an edge preserving property but sometimes it gives undesirable blurred effect. We prove the existence and uniqueness theorem of our proposed model. The results of our model using explicit numerical schemes are compared with other known image restoration models.

Keywords:Second order PDE's, image denoising, nonlinear diffusion.

AMS Subject Classification: 35K51, 35K55, 35K99.

1. INTRODUCTION

Image denoising is a fundamental problem in both image processing and computer vision with numerous applications. The total variation models [7, 19, 20] and anisotropic diffusion models [6, 18, 22, 23, 24] have been studied as a useful tool to the problem of image denoising and image reconstruction. These partial differential equation based image enhancement techniques have been able to achieve a good edge preservation. An image can be interpreted as a real function defined on Ω, a bounded and open domain of $\mathbb{R}^2$ (for simplicity we will assume Ω to be the square domain henceforth). Formation of a noisy image is typically modeled as

$$u_0(x,y) = u(x,y) + n(x,y),$$

where $u(x,y)$ denote the desired clean image, $u_0(x,y)$ denote the pixel values of a noisy image for $x, y \in \Omega, \Omega$ is a bounded open subset of $\mathbb{R}^2$ and $n(x,y)$ is additive white noise assumed to be close to Gaussian. The values $n(i,j)$ of n at the pixels (i,j) are independent random variables, each with a Gaussian distribution of zero mean and variance σ^2.

We propose the following second order - version of the nonlinear diffusion model which is a synthesis of ideas from Catté et al. [6]. Our model is given by:

$$\begin{aligned} &\frac{\partial u}{\partial t} = \nabla\cdot(g_1(|\nabla G_\sigma * u|)\nabla u) + \nabla\cdot(g_2(|\nabla G_\sigma * u|)\nabla u) - \lambda(u - u_0) \text{ on } \Omega \times (0,T), \\ &u(x,0) = u_0(x) \text{ in } \Omega, \\ &\frac{\partial u}{\partial \vec{n}} = 0 \text{ on } \partial\Omega \times (0,T), \end{aligned} \tag{1}$$

where g_1 and g_2 are decreasing function tending to zero at infinity with $g_i(0) = M_i > 0$ and $t \to g_i(\sqrt{t})$ is smooth, and $G_\sigma(x)$ is the Gaussian kernel, namely,

$$G_\sigma(x) = \frac{1}{2\pi\sigma^2} e^{\frac{-|x|^2}{2\sigma^2}}. \tag{2}$$

Department of Mathematics Aligarh Muslim University, Aligarh-202002, India

e-mail: skykumar87@gmail.com
e-mail: ahmad_kalimuddin@yahoo.co.in
Manuscript received xx.

Paper-5

[NPSC] Submission Acknowledgement

3 messages

Michael P. Bekakos <mbekakos@ee.duth.gr> Wed, Feb 24, 2016 at 12:37 PM
To: Santosh Kumar <skykumar87@gmail.com>

Santosh Kumar:

Thank you for submitting the manuscript, "An efficient PDE-based nonlinear anisotropic diffusion model for image denoising" to Neural, Parallel, and Scientific Computations. With the online journal management system that we are using, you will be able to track its progress through the editorial process by logging in to the journal web site:

Manuscript URL:
http://www.dynamicpublishers.org/journals/index.php/NPSC/author/submission/295
Username: sk_rs-kumar

If you have any questions, please contact me. Thank you for considering this journal as a venue for your work.

Michael P. Bekakos
Neural, Parallel, and Scientific Computations

Neural, Parallel, and Scientific Computations
http://dynamicpublishers.org/journals/index.php/NPSC

---

santosh kumar <skykumar87@gmail.com> Wed, Feb 24, 2016 at 3:08 PM
To: Sarfaraz Mohammad <sarfarazm820@gmail.com>

[Quoted text hidden]

---

santosh kumar <skykumar87@gmail.com> Fri, Mar 4, 2016 at 10:45 AM
To: ahmad_kalimuddin@yahoo.co.in

[Quoted text hidden]

Neural, Parallel, and Scientific Computations

Active Submissions

ACTIVE ARCHIVE

ID	MM-DD SUBMIT	SEC	AUTHORS	TITLE	STATUS
295	02-24	ART	Kumar	AN EFFICIENT PDE-BASED NONLINEAR ANISOTROPIC DIFFUSION...	Awaiting assignment

1 - 1 of 1 Items

Start a New Submission

CLICK HERE to go to step one of the five-step submission process.

Refbacks

ALL NEW PUBLISHED IGNORED

DATE ADDED	HITS	URL	TITLE	STATUS	ACTION

There are currently no refbacks.

ISSN: 1061-5369

An efficient PDE-based nonlinear anisotropic diffusion model for image denoising

Santosh Kumar, Mohd. Sarfaraz and M.K. Ahmad
Department of Mathematics
Aligarh Muslim University
Aligarh, 202002, India
E-mails: skykumar87@gmail, sarfarazm820@gmail.com, ahmad_kalimuddin@yahoo.co.in

Abstract

In this paper, we propose a new nonlinear anisotropic diffusion model for image denoising. The main idea is to apply a priori smoothness on the solution image. We present proof of the viscosity solution of our model. The results of our model using explicit numerical schemes are compared with other known image denoising models.
Keywords: Image denoising, nonlinear diffusion

1 Introduction

The nonlinear diffusion techniques and PDE-based variational models are very popular in image denoising and restoration. The nonlinear diffusion method for image denoising and edge detection was first introduced by Perona and Malik [13]. This method is based on a diffusion process governed by a partial differential equation (PDE), where diffusion amount depends on the gradient of images.

Mathematically, $u_0 : \Omega \to \mathbb{R}$ represents a noisy version of a true image, and it is obtained by the following imaging process

$$u_0(x) = u(x) + n(x), \tag{1.1}$$

where $u(x)$ denotes the desired clean image, $u_0(x)$ denotes the pixel values of a noisy image for $x \in \Omega$, $\Omega \subset \mathbb{R}^2$ is a bounded domain, usually a rectangle and $n(x)$ is additive white noise assumed to be close to Gaussian. The values $n(i, j)$ of n at the pixels (i, j) are independent random variables, each with a Gaussian distribution of zero mean and variance σ^2.

Paper-6

Gmail santosh kumar <skykumar87@gmail.com>

Thank you for your approval

1 message

ISTT - Editorial Office <em@editorialmanager.com> Fri, Apr 1, 2016 at 1:26 PM
Reply-To: ISTT - Editorial Office <neeraj.poduval@springer.com>
To: Santosh Kumar <skykumar87@gmail.com>

Dear Dr. Kumar,

Thank you for approving the changes and returning your submission entitled "Denoising method based on wavelet coefficients via diffusion equation".

You will be able to check on the progress of your paper by logging on to Editorial Manager as an author. The URL is http://istt.edmgr.com/.

Thank you for submitting your work to this journal.

Kind regards,
Editorial Office
Iranian Journal of Science and Technology, Transactions A: Science

Iranian Journal of Science and Technology, Transactions A: Science

Denoising method based on wavelet coefficients via diffusion equation

--Manuscript Draft--

Manuscript Number:	ISTT-D-16-00179
Full Title:	Denoising method based on wavelet coefficients via diffusion equation
Article Type:	Original Paper
Funding Information:	
Abstract:	Image denoising is still a crucial problem for image processing community as well as mathematicians alike. This paper propose a denoising technique based on wavelet coefficients via diffusion equation. The present work compares different denoising process by diffusion models. The data is initially subjected to synthetic noisy data with various levels of standard deviation ${\sigma}^2$. To quantify the results, we use peak signal to noise ratio (PSNR) as metric.
Corresponding Author:	Santosh Kumar, Ph.D Aligarh Muslim University, Aligarh Aligarh, Uttar-Pradesh INDIA
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Aligarh Muslim University, Aligarh
Corresponding Author's Secondary Institution:	
First Author:	Santosh Kumar, Ph.D
First Author Secondary Information:	
Order of Authors:	Santosh Kumar, Ph.D
	Mohd. Sarfaraz, M.Sc.
	Md. Kalimuddin Ahmad, Ph.D
Order of Authors Secondary Information:	
Author Comments:	
Suggested Reviewers:	Khalil Ahmad, Ph.D Professor, Jamia Millia Islamia, New Delhi, India khalil_ahmad49@yahoo.com Working in the same field.
	Rathish Kumar, Ph.D Professor, Indian Institute of Technology Kanpur bvrk@iitk.ac.in Working in same field.
	S. Sundar, Ph.D Professor, Indian Institute of Technology Madras slnt@iitm.ac.in Working in the same field.
	Dhanesh Patel, Ph.D Professor, Maharaja Sayajirao University of Baroda pdhanesh@yahoo.com Working in the same field.

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Blinded Manuscript Click here to view linked References

Noname manuscript No.
(will be inserted by the editor)

Denoising method based on wavelet coefficients via diffusion equation

Received: date / Accepted: date

Abstract Image denoising is still a crucial problem for image processing community as well as mathematicians alike. This paper propose a denoising technique based on wavelet coefficients via diffusion equation. The present work compares different denoising process by diffusion models. The data is initially subjected to synthetic noisy data with various levels of standard deviation σ^2. To quantify the results, we use peak signal to noise ratio (PSNR) as metric.

Keywords Discrete wavelet transform · image denosing · anisotropic diffusion

1 Introduction

Interest in digital image processing methods stems from two principal application areas:

(a) Improvement of pictorial information for human interpretation
(b) Processing of image data for storage, transmission, and representation for autonomous machine perception.

The crucial point in these approaches is to distinguish between important features that should be kept or even enhanced. Mathematically different ways have been used to model how a smooth image should look like, i.e., a certain smoothness can be formalised in terms of differentiability orders and small modulus of derivatives. This idea leads to regularisation methods and related partial differential equations. Another kind of smoothness assumption is to discard certain wavelet coefficients, leading to the popular wavelet shrinkage method.

For denoising case, the underlying model for the image degradation is an additive noise model. Thus in discrete setting, we have three images $u_0, u, n \in$